MINING ASSOCIATION RULES BETWEEN CREDITS IN THE LEADERSHIP IN ENERGY AND ENVIRONMENTAL DESIGN FOR NEW CONSTRUCTION (LEED-NC) GREEN BUILDING ASSESSMENT SYSTEM

THESIS

Benjamin J. Thomas, Second Lieutenant, USAF

AFIT/GEM/ENV/08-M19

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GEM/ENV/08-M19

MINING ASSOCIATION RULES BETWEEN CREDITS IN THE LEADERSHIP IN ENERGY AND ENVIRONMENTAL DESIGN FOR NEW CONSTRUCTION (LEED-NC) GREEN BUILDING ASSESSMENT SYSTEM

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Engineering Management

Benjamin J. Thomas, BS

Second Lieutenant, USAF

March 2008

AFIT/GEM/ENV/08-M19

MINING ASSOCIATION RULES BETWEEN CREDITS IN THE LEADERSHIP IN ENERGY AND ENVIRONMENTAL DESIGN FOR NEW CONSTRUCTION (LEED-NC) GREEN BUILDING ASSESSMENT SYSTEM

Benjamin J. Thomas, BS

Second Lieutenant, USAF

Approved:

// Signed //______________________________	_3-11-08_
Alfred E. Thal Jr. (Chair)	Date
// Signed //______________________________	_3-11-08_
Gilbert Peterson (Member)	Date
// Signed //______________________________	_3-11-08_
Nadja Turek, Capt, USAF (Member)	Date

AFIT/GEM/ENV/08-M19

Abstract

The Leadership in Energy and Environmental Design (LEED) Building Assessment System is a performance-based tool for determining the environmental impact of a facility from the whole-building perspective. Taking this vision into account, the individual credits that comprise LEED are designed to reward design teams for employing sustainable design strategies that reduce the total environmental impact across several sustainability issues. This study analyzed projects that have been certified in LEED for New Construction (LEED-NC) versions 2.0 and 2.1. Data on the credits achieved by the projects were mined using the Apriori algorithm which produced 641 association rules. These results were then subjectively reduced to the 24 most synergistic credit combinations and were subsequently identified as credit bundles. This study provided insight into credit interplay and its effect on high-scoring sustainable design strategies. Additionally, it shows that no one strategy is systematically employed by sustainable design professionals in the pursuit of LEED certification. This research lays the foundation for the application of data mining techniques to future LEED data sets. Finally, the revealed credit bundles support the assertion that LEED is a tool that rewards whole-building design and reinforces the perception that integrated design teams are a critical element in successful LEED project delivery.

Acknowledgments

I would like to express my sincere appreciation to the members of my committee, especially to my advisor Dr. Al Thal for his patience and guidance. Also, I owe a huge debt of gratitude to the Director of Technical Development for LEED, for being bold enough to listen to the ideas of a rogue graduate student he had never heard of. There is also his staff, without whom the support, data, and relevancy of this research would not exist. Their collaboration was integral in the deliverance of this work. Of course, my most profound thanks belongs to my wife whose constant love and persistent encouragement ensured the completion of this work despite my nature to procrastinate.

Benjamin J. Thomas

Table of Contents

Page

Page

Page

List of Figures

Page

List of Tables

Page

MINING ASSOCIATION RULES BETWEEN CREDITS IN THE LEADERSHIP IN ENERGY AND ENVIRONMENTAL DESIGN FOR NEW CONSTRUCTION (LEED-NC) GREEN BUILDING ASSESSMENT SYSTEM

I. Introduction

The construction industry can trace its heritage back to the Egyptian pyramids and perhaps even further through the history of human existence. Throughout that history, the built environment has expanded steadily and the industry now finds itself in a new paradigm. The construction necessary to support the exponential growth of the human population consumes vast resources and produces such large quantities of waste that some are concerned the strain will soon overcome earth's capacity (Hardin, 1968). The construction industry's response to this concern is “sustainable development,” sometimes referred to as “green building.” In 1987, the United Nations convened the World Commission on the Environment and Development which produced a report that defined sustainable development as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED, 1987). Known as the Brundtland Report, so named after the commission's chairperson, it is commonly viewed as the first serious global discussion about sustainable development.

Background

Approximately 65% of electricity consumption and nearly 36% of total energy use in the United States is accounted for in buildings; additionally 30% of greenhouse gas emissions, 30% of raw materials, 136 million tons of waste annually, and 12% of potable water are attributed to commercial buildings (USGBC, 2008). The technology to mitigate these daunting impacts already exists and is constantly improving. Development in building energy systems, like heating and air conditioning equipment, has brought about increasingly efficient performance. Additionally, research into new materials and recycling of materials for the construction industry is helping to alleviate the strain on raw materials and waste streams.

Unfortunately, there is no one single product or "green" solution that can make the entire built environment sustainable by itself. Green materials, methods, and equipment are only components of complex modern facilities. The existence, operation, and maintenance of the whole building are the cause of the environmental impacts. For this reason, the whole building must be taken into consideration in order to properly determine those impacts. Based on the severity of the environmental impacts and the efforts taken to mitigate them, building owners can begin to assess the sustainability of their construction projects.

For the purpose of rating the level of sustainability a building achieves, several assessment tools have been developed. Britain established the Building Research Establishment Environmental Assessment Method (BREEAM), Australia uses the Green

Star program, and Japan has recently developed the Comprehensive Assessment System for Building Environmental Efficiency (CASBEE). In the United States, the Leadership in Energy and Environmental Design (LEED) Green Building Rating System, a product of the U.S. Green Building Council (USGBC), has emerged as the accepted standard (Kibert, 2005).

LEED consists of a series of credits that are awarded based on the performance of a building, not on the methods used. The credits are available in six categories: Sustainable Sites (SS), Water Efficiency (WE), Energy and Atmosphere (EA), Materials and Resources (MR), Indoor Environmental Quality (EQ), and Innovation and Design Process (ID). Depending on the number of credits accumulated, LEED certifies buildings at four different levels: certified, silver, gold, and platinum (USGBC, 2005). In 1998, the LEED standard for New Construction and Major Renovations was released, known as LEED-NC version 1.0. Since 2000, the year LEED-NC was updated to version 2.0, several other LEED products have been released including standards for Existing Buildings (EB), Commercial Interiors (CI), Homes (H), and Core and Shell (CS) (Kibert, 2005). New variants and updated versions of current LEED standards are constantly under development; as of the writing of this report, the latest version of LEED-NC is version 2.2 and will stand as the reference standard for the purposes of this study (USGBC, 2007).

From the USGBC's inception, government organizations have played a significant role in the promotion of sustainable development and support of LEED

products. In early 2006, the heads of 21 federal organizations, including the Department of Defense, signed a memorandum of understanding pledging a commitment to leadership in design, construction, and operation of high performance sustainable buildings (MOU, 2006). Although the memorandum is not enforceable, it is an important development considering the fact that the federal government accounts for 455,000 buildings with over 3.0 billion square feet not to mention hundreds of millions of leased square footage (OFEE, 2006). Department of Defense assets account for more than two-thirds of federal buildings, so its construction policy is crucial to the proliferation of sustainable practices (OFEE, 2006).

Currently, both the Army and the Navy mandate LEED certification. The Army Corps of Engineers requires silver certification for all new military construction (MILCON) projects and is taking an active role in the LEED for homes and the LEED for neighborhood development pilot programs (Army Memo, 2006). The Navy requires all construction to conform to certifiable LEED standards for their new construction. Larger projects are required to actually register for the certification, and high-visibility projects must achieve at least silver (NAVFAC, 2005). Until recently, the Air Force employed the most conservative sustainability policy, requiring simply that its new facilities be certifiable under LEED standards, without mandating actual registration (AF Memo, 2001). On July 31st 2007, the Air Force stepped up their policy and mandated that by 2009 all new construction projects will be designed to silver-certifiable levels with 5% actually registering for certification with 10% in 2010 and thereafter. All other projects

will be reviewed internally by LEED Accredited Professionals to ensure that standards are being met (AF Memo, 2007).

Problem Statement

One of the most common fears among building owners and developers is that green buildings cost more to construct. This issue has been studied thoroughly and most studies have concluded that the marginal cost premium of LEED certified buildings is anywhere from 0 to 9% (Matthiessen and Morris, 2004; GSA, 2005) with an average of about 2% (Katz, 2003). Matthiessen and Morris (2004) even go as far as to conclude that the marginal cost premium associated with LEED-seeking buildings falls within the variability of equivalent non-LEED-seeking construction. A common strategy employed to reduce first-costs in LEED construction is to select the credits early and aim for credits that are synergistic (GSA, 2005). Consequently, there are many sustainable design strategies that achieve several credits at once; these are often referred to as credit synergies or credit bundles.

Although credit selection is critical to the success of LEED certified construction projects, no studies have been conducted to directly address the aggregate trends and relationships between particular credits. There are over 800 buildings certified under LEED-NC version 2.0 and later, with thousands of projects registered and pending certification. Taken together, these projects constitute thousands of decisions about credit appropriateness, cost-benefit ratios, and feasibility by owners, architects, designers,

and other construction professionals from around the country. If there are strong associations between credits, that information can be applied to expedite the credit selection process or can reveal common credit bundles employed by the industry. This knowledge and the process used to discover it may give the USGBC new insight into the LEED-NC product that could be applied to later versions of the LEED system or addendums to the Reference Guide.

Research Questions

By identifying credits with strong associations, credit bundles that are commonly achieved can be identified. These credit selection biases may reveal interesting implications for current sustainability practitioners and for the future development of LEED. The most important revelation is whether or not the synergies between certain credits can be used by the USGBC to streamline the credit review process. Therefore, the following questions are addressed in this research:

1. What are the most interesting credit bundles?
2. Are there any actionable implications of the revealed credit synergies?
3. Can the USGBC use association rules to streamline the credit review process?

Methodology

A form of data mining known as association rule mining was used to reveal the associations between credits in LEED. Data mining is the automated or convenient

extraction of patterns representing knowledge implicitly stored or captured in large databases or data warehouses (Han and Kambler, 2006). Association rules are presented in a Boolean format commonly expressed as “if-then" statements. An example of an association rule for LEED credits is “*if* a project achieves Sustainable Sites Credit 6.2, *then* the project is also likely to achieve Sustainable Sites Credit 7.1.” This rule asserts that SS 6.2 predicts SS 7.1 and, therefore, the two credits form a credit bundle. The reason for employing data mining as opposed to more traditional statistical techniques is that association rule mining is a form of undirected learning (Barry and Linoff, 2004). This means that instead of searching for a correlation based on some hypothesized relationship, the association rule mining algorithm searches the entire data set for associations and simply reports the strongest ones.

Limitations

There are several limitations that accompany the circumstances and methods employed in this data mining exercise. First, the only credit bundles that surface are the ones popular enough with the industry to be selected by multiple project teams. There may be many credit synergies inherent in the LEED-NC system that simply are not popular enough with owners and consultants to emerge in the analysis. Secondly, not all credit synergies lead to cost savings. There are several credit bundles achieved by very expensive design strategies; for example, an intensive green roof can add anywhere from $30 to $100 per square foot to a building’s cost. However, one can assume that most

building owners are concerned with minimizing construction costs; therefore, most of the popular credit bundles revealed in this study are those that reflect cost savings. That being said, there is still the possibility that the credit bundles that emerge have less to do with cost savings and more to do with perceived cost savings and credit bundling preferences of the leading green consultants. In other words, some of the credit bundles revealed in this study may be self-fulfilling prophesies of perceived synergies rather than actual synergies inherent in the LEED system.

There are also limitations introduced by the methodology chosen for this analysis. Although association rules are based in probabilistic statistics, they are not the strongest method for supporting a proposed relationship. Causation with a correlation of statistical significance and the absence of alternative explanations is the accepted scientific method for supporting a hypothesized relationship. In defense of the chosen methodology, empirically proving the relationships between the revealed credits is not the primary purpose of the analysis. The primary objective, like most other applications of data mining, is to fill in the gaps in intuition. For this task, data mining is the best practice available.

Another key limitation in analyzing LEED project data is that green construction is an emerging industry still in its growth phase. LEED for New Construction and Major Renovations is the USGBC's oldest and most popular product. As of October 2007, there were over 800 certified projects and more than 5,000 projects registered and awaiting certification in the LEED-NC system (USGBC, 2008). That represents nearly 70% of the

active projects for all seven LEED products. As impressive as these numbers are, it is important to note that project registrations have been growing exponentially each year. Even with nearly a thousand data points, one cannot consider the green construction movement to be in a steady state. Coupled with that limitation is the fact that the LEED standards are themselves in a constant state of flux. The version of LEED-NC referenced by this study, version 2.2, is the fourth iteration in just 7 years. There have already been several seminars on the topic of LEED's next iteration, version 3.0 (USGBC, 2008). This continuous improvement is necessary for the USGBC to maintain its goal of market transformation by targeting "the leading 25% of best practice shown by early adopters" (USGBC, 2006). The technologies employed to deliver sustainability to that "leading 25%" of the industry are at the cutting-edge of the construction industry. For that reason, there is no foreseeable point in time when LEED project data will be at a "steady state." For this reason, credit synergies identified by this study should not be considered for application to projects certified under future versions of LEED.

Chapter Overview

The remainder of this thesis document will present the research conducted in order to identify credit bundles in the LEED-NC building assessment system. It begins in Chapter II with a review of the literature. This includes sections on LEED and its structure, as well as data mining and case studies applying similar methodologies. In Chapter III, the methodology for this research is explained in full detail. The algorithm

used to mine the association rules and the process employed in the analysis are discussed in depth. In Chapter IV, the results are presented along with discussion to answer the research questions. Finally, Chapter V presents the conclusions reached during the research. This includes recommendations for future research and discussion on its impact on the Air Force, the USGBC, and the greater sustainable development community.

II. Literature Review

In order to identify credit bundles and gain insight into the associations between credits, it is critical to first understand the structure of the Leadership in Energy and Environmental Design (LEED) green building assessment system. Then it becomes necessary to investigate the best practices and established processes in data mining techniques to ensure that the chosen approach is appropriate. A section expounding on the concepts in association rule mining followed by a section defining the key metrics is provided to support further discussion on association rule mining practice. Before establishing a methodology, it is best to review other case studies that provide examples of successful data mining application to similar topics.

LEED Building Assessment System

The earliest attempt at forming a standard for assessing green buildings in the United States was taken in 1994 by the American Society for Testing and Materials (ASTM). The first iteration was unsuccessful, but served as the foundation for LEED which was created by the United States Green Building Council (USGBC). The first LEED product was released in 1998 as a pilot version that applied only to new construction. There were only 18 projects in that pilot study. In 2000, the USGBC released LEED version 2.0 in a form recognizable by today's green building industry with 69 available points and four levels of certification (Kibert, 2005).

Structure of LEED

Through versions 2.0, 2.1, and 2.2, LEED has maintained the same basic structure: five core categories and one bonus category. Those categories are Sustainable Sites (SS), Water Efficiency (WE), Energy and Atmosphere (EA), Materials and Resources (MR), Indoor Environmental Quality (EQ), and Innovation and Design Process (ID) (USGBC, 2005). The SS category contains credits aimed at minimizing the project's impact on the local environment. It spans from issues of restoring natural habitat to credits for discouraging internal-combustion-powered automobile transportation. The WE category provides credits for reducing water use which is an important environmental issue in densely populated cities and areas west of the Rocky Mountains. The EA category aims to reduce the impact of energy usage and generation on the atmosphere. It accomplishes this by providing credits for increased efficiency of building mechanical systems and credits that encourage the use of renewable energy and eco-friendly refrigerants. The MR category attempts to minimize the effect construction materials have on the environment. It provides credits for several materials issues including demolition waste management and encouraging the use of recycled materials. The last of the core categories, EQ credits are focused on the health and comfort of building occupants. This includes issues regarding indoor air quality, natural lighting, ventilation, and the elimination of toxic substances.

The final category in LEED-NC is the bonus ID category. There are four "innovation in design" points that are awarded for exemplary performance in some of the

core category credits or for successfully applying a sustainable strategy that is not covered by an existing credit. Additionally, a single point is available for having a LEED Accredited Professional (AP) on the project team. Accreditation is available to any interested persons through the USGBC and consists of passing a 2-hour proctored exam for a reasonable fee (USGBC, 2005).

The relative importance of each category is a function of how many points the USGBC has made available in that category. For example, the Energy and Atmosphere category is the largest with up to 17 possible points, while the water efficiency category is the smallest with only 5 available points. Some of the categories have prerequisites that are often listed with the credits but do not count towards the tallying of points for determining the level of certification achieved. A comprehensive list of all 69 points from the LEED for New Construction (LEED-NC) version 2.2 system is available in Appendix A.

A common source of confusion when discussing the LEED system is the difference between credits and points. There are 34 credits in LEED, some of which are worth multiple points, adding up to the 69 available points in the LEED-NC system. Each credit represents a specific sustainability issue. The multi-point credits can offer different points representing unique approaches to addressing the credit's sustainability issue; alternatively, they can offer several points representing increments of the same approach. For example, in the Sustainable Sites (SS) category, credit 4 is worth up to

four points for four unique and different approaches to encouraging use of alternative transportation as shown in Table 1.

Table 1. Sustainable Sites (SS) category from LEED-NC, Version 2.2

Credit	Description
SS 1	**Site Selection**
SS 2	**Development Density & Community Connectivity**
SS 3	**Brownfield Redevelopment**
SS 4.1	**Alternative Transportation**, Public Transportation Access
SS 4.2	**Alternative Transportation**, Bicycle Storage & Changing Rooms
SS 4.3	**Alternative Transportation**, Low-Emitting & Fuel-Efficient Vehicles
SS 4.4	**Alternative Transportation**, Parking Capacity
SS 5.1	**Site Development,** Protect or Restore Habitat
SS 5.2	**Site Development,** Maximize Open Space
SS 6.1	**Stormwater Design,** Quantity Control
SS 6.2	**Stormwater Design,** Quality Control
SS 7.1	**Heat Island Effect,** Non-Roof
SS 7.2	**Heat Island Effect,** Roof
SS 8	**Light Pollution Reduction**

However, in the Energy and Atmosphere (EA) category, credit 2 is worth up to three points. The first point is awarded for installing renewable energy sources on the project site that provide at least 2.5% of the project's average daily energy needs. The other two

are awarded for increasing that threshold to 7.5% and 12.5%, respectively. The EA category is displayed in Table 2.

Table 2. Energy and Atmosphere (EA) category from LEED-NC, Version 2.2

Credit	Description
EA 1.1	**Optimize Energy Performance,** 10.5% New or 3.5% Existing
EA 1.2	**Optimize Energy Performance,** 14% New or 7% Existing
EA 1.3	**Optimize Energy Performance**, 17.5% New or 10.5% Existing
EA 1.4	**Optimize Energy Performance,** 21% New or 14% Existing
EA 1.5	**Optimize Energy Performance,** 24.5% New or 17.5% Existing
EA 1.6	**Optimize Energy Performance,** 28% New or 21% Existing
EA 1.7	**Optimize Energy Performance,** 31.5% New or 24.5% Existing
EA 1.8	**Optimize Energy Performance,** 35% New or 28% Existing
EA 1.9	**Optimize Energy Performance,** 38.5% New or 31.5% Existing
EA 1.10	**Optimize Energy Performance,** 42% New or 35% Existing
EA 2.1	**On-Site Renewable Energy,** 2.5%
EA 2.2	**On-Site Renewable Energy,** 7.5%
EA 2.3	**On-Site Renewable Energy,** 12.5%
EA 3	**Enhanced Commissioning**
EA 4	**Enhanced Refrigerant Management**
EA 5	**Measurement & Verification**
EA 6	**Green Power**

Credit Bundles

For the purposes of this study, a credit bundle is defined as a group of two or more points or credits from LEED-NC that are synergistic. The concept of credit synergy was not created by consultants and design professionals; it was intentionally built into the LEED system because of the nature of sustainable development. The environmental impact of a vertical construction project is a result of the whole building, not just the sub-systems that comprise it. For this reason, LEED was conceived with a performance-based structure that requires project teams to apply an integrated design process, thereby helping break down the historic barriers between the various construction disciplines. This performance-based approach allows design strategies that address multiple sustainability issues to thrive under LEED where the only metric that matters is how many points one accumulates, regardless of how they are achieved. The USGBC has included several sub-sections into the Reference Guide regarding "credit synergies" for certain credit descriptions (USGBC, 2005).

In 2004, the U.S. General Services Administration (GSA) released a comprehensive study of the cost of applying LEED to typical GSA construction projects. The purpose of their study was to estimate the cost of developing green federal facilities using LEED-NC Version 2.1. Their methodology consisted of applying several cost estimates for two hypothetical construction projects: new construction of a mid-rise federal courthouse, and the modernization of an existing mid-rise federal office building. In the final report, an entire section was dedicated to "synergistic" credits because it was

the authors' belief that "identifying and exploiting the synergies among LEED credits is often a key step in achieving successful, cost-effective green projects" (GSA, 2004). The GSA's report specifically identifies five unique credit bundles. A list of those bundles and a brief description of the synergy between the credits is provided in Table 3. The authors of the GSA cost study are quick to point out that their list of credit synergies is not comprehensive (GSA, 2004).

Table 3. Credit Bundles Identified by the GSA Cost Study (GSA, 2004)

Credits in the Bundle	Reason for Synergy
SS-5.1, SS-6.1, WE-1.1/1.2	Restoring natural area improves permeability and reduces irrigation.
SS-6.1, SS-7.2	Green roofs reduce storm water run-off and heat island effect.
EA-1, EQ-1	CO2 Monitors are a critical component in most high-performance HVAC Systems.
MR-7, EQ-4.4	Lack of certified woods with added urea formaldehyde resins
MR-4.2, MR-5.2	Local gypsum recycler available in the area

Data Mining

Data mining is the analysis of observational data sets to "find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner" (Hand, 2001). An important part of this definition is that the data is observational. While data that is collected exclusively for the purposes of analysis

is experimental data, observational data is collected for other reasons (Hand, 2001). For this reason, most data mining initiatives are examples of "secondary" data analysis (Hand, 2001). Since the USGBC stores project-specific credit information as part of their electronic records for auditing and continuity purposes, and not necessarily for data analysis, historical LEED credit achievement data is considered observational.

The reference to the understandability and usefulness of the data mining results is also very important in this definition of data mining. It implies that data mining is primarily an applied field. The knowledge discovered through the application of data mining is used to advance or fill-in the gaps in existing intuition gained through experience in the subject area. The purpose of data mining is not to provide the strongest possible empirical evidence for some hypothesis. Rather, the result of the data mining study can be viewed as a hypothesis. Researchers who are concerned with providing the most possible empirical support for some hypothesis should apply the scientific method. Data mining is best applied in subject areas that rely mostly on intuition, like marketing, sales, and customer relationship management (Berry and Linoff, 2004). Commonly, data mining is considered a part of the broader concept of knowledge discovery in databases (KDD) (Fayyad, 1996).

Regardless of the label one uses, data mining is an interdisciplinary field that encompasses statistics, database technology, machine learning, artificial intelligence, and high-performance computing to name a few (Han and Kambler, 2006). It can be viewed as a result of a natural progression in information technology. Thanks to the information

age and advances in data warehousing, organizations have been automating process and generating vast pools of data. Conventional thinking is that there is something to be learned from all this data, and data mining is the result of attempts at discovering knowledge in that data (Han and Kambler, 2006).

Data mining algorithms are simply automated techniques for analyzing large databases that are simply too vast or too complex for traditional data analysis methods. However, they are not intended to replace human intelligence, but rather to enhance existing intuition about the topic at hand (Larose, 2005). The proper application of data mining requires an analyst that is well-versed in the business that generated the data and understands the underlying principles of the algorithm being applied. Data mining is very easy to do poorly because the algorithms are designed to produce results and will rarely fail to do so (Larose, 2005). It is up to the human element to determine whether or not the output is relevant or trivial and whether it is actionable or just an anomaly (Berry and Linoff, 2004).

The Data Mining Process

Most texts on data mining provide a recommended process for data mining initiatives in an early chapter (Berry and Linoff, 2004; Hand, 2001; Fayyad, et al., 1996). The reason it is so critical to use a methodical approach to data mining is because haphazardly applying data mining algorithms to a data set can lead to inferring knowledge that is either not true or true but not useful (Berry and Linoff, 2004). It is

very possible to produce an entire list of patterns and rules that are trivial. Like any form of data analysis, it is up to an intelligent human element to determine whether or not the results generated are of value (Larose, 2005).

With these concerns in mind, a group of data mining practitioners from DaimlerChrysler Inc., SPSS Inc., and NCR Inc. developed the Cross-Industry Standard Process for Data Mining (CRISP-DM) in 1996 (Larose, 2005). Although CRISP-DM was developed by practitioners and not academics, it is the accepted methodology applied by the majority of data mining professionals (KDNuggets.com, 2007). This is most likely because the method is developed and mutually agreed upon by a consortium of data miners called the CRISP-DM Special Interest Group. This group consists of a wide range of practitioners, including data warehouse vendors and management consultancies, with a vested interest in data mining applications (Chapman, et al., 2000).

The CRISP-DM method consists of six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Chapman, et al., 2000). The reference model for a typical data mining project is shown in Figure 1. The shape of the reference model is intended to convey the cyclic nature of data mining initiatives. It should be pointed out that deployment is not necessarily the end of a data mining initiative. Often, the deployment of one solution leads to more focused business questions that develop into another data mining cycle. The arrows between the phases reveal that data mining is not necessarily a sequential endeavor. Depending on the outcome of each phase, prior phases may have to be revisited or even redone. The arrows

are not intended to represent all the relationships among the six phases, only the most common paths (Chapman, et al., 2000).

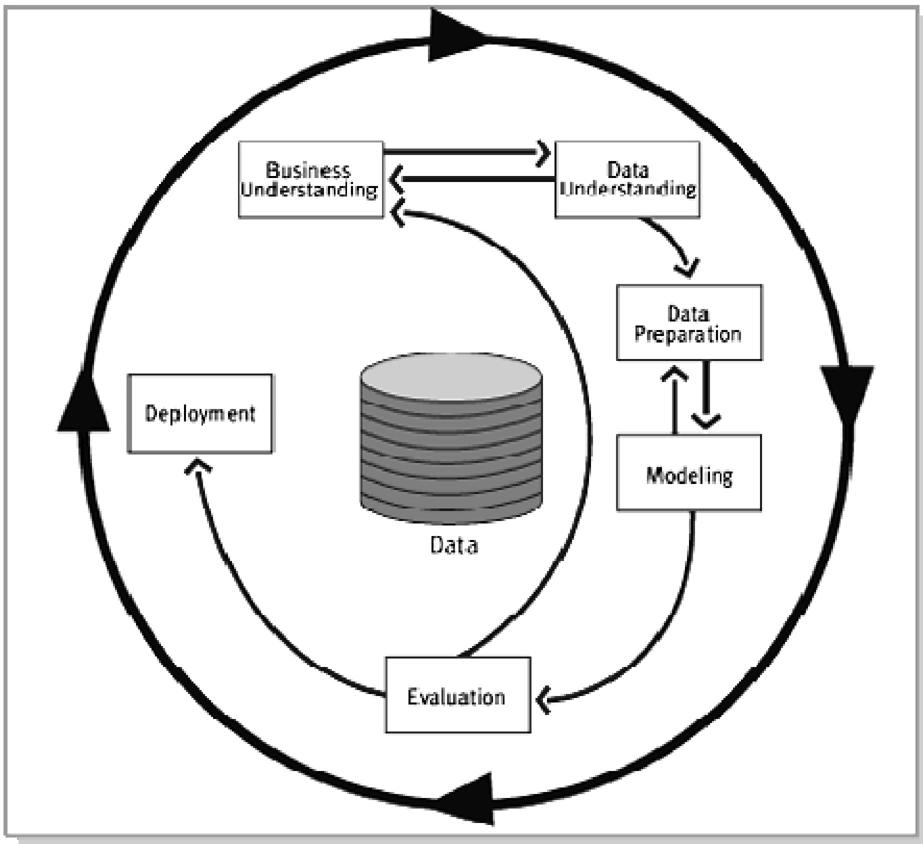

Figure 1. Phases of the CRISP-DM reference model (Chapman, et al., 2000)

The initial business understanding phase focuses on gaining a perspective of the business and translating that into a data mining problem. The data understanding phase

includes gathering the data, typically from observational data warehouses, and activities aimed at familiarization and quality control of the data. The modeling phase is when the data miner will select a data mining technique, set of techniques, or develop their own data mining technique. With the model or models built, the evaluation phase is when the data mining results are scrutinized to assess whether or not the results actually answer the business problem established in the project outset. Finally, the results of the data mining initiative are deployed into a business solution (Chapman, et al., 2000). Deployment initiatives can range from an ongoing real-time data mining solution that is integrated into existing business processes to simply a report explaining the new knowledge gained. Typically, the deployment phase is carried out by the business process owner and not by the data mining practitioner (Chapman, et al., 2000). The general tasks and outputs from each phase are highlighted in Figure 2.

Business Understanding

Determine Business Objectives
Background
Business Objectives
Business Success Criteria

Assess Situation
Inventory of Resources
Requirements, Assumptions, and Constraints
Risks and Contingencies
Terminology
Costs and Benefits

Determine Data Mining Goals
Data Mining Goals
Data Mining Success Criteria

Produce Project Plan
Project Plan
Initial Assessment of Tools and Techniques

Data Understanding

Collect Initial Data
Initial Data Collection Report

Describe Data
Data Description Report

Explore Data
Data Exploration Report

Verify Data Quality
Data Quality Report

Data Preparation

Select Data
Rationale for Inclusion/Exclusion

Clean Data
Data Cleaning Report

Construct Data
Derived Attributes
Generated Records

Integrate Data
Merged Data

Format Data
Reformatted Data

Dataset
Dataset Description

Modeling

Select Modeling Techniques
Modeling Technique
Modeling Assumptions

Generate Test Design
Test Design

Build Model
Parameter Settings
Models
Model Descriptions

Assess Model
Model Assessment
Revised Parameter Settings

Evaluation

Evaluate Results
Assessment of Data Mining Results w.r.t. Business Success Criteria
Approved Models

Review Process
Review of Process

Determine Next Steps
List of Possible Actions
Decision

Deployment

Plan Deployment
Deployment Plan

Plan Monitoring and Maintenance
Monitoring and Maintenance Plan

Produce Final Report
Final Report
Final Presentation

Review Project
Experience Documentation

Figure 2. Generic tasks (bold) and outputs (italic) of the CRISP-DM reference model (Chapman, et al., 2000)

Although the make-up and sequence of the phases and their related sub-tasks is fairly intuitive, further description is available in the "CRISP-DM Step-by-step Data Mining Guide" available on the website: www.crisp-dm.org. It is not necessary to replicate the entire process as outlined. However, the CRISP-DM reference model does provide a very intensive and generally-accepted approach to data mining practice.

Data Mining Techniques

Although data mining is an emerging and continuously evolving field, there are five recognized umbrella tasks under which data mining projects are labeled (Hand, 2001). The first is exploratory data analysis (EDA). As the name implies, the purpose of an EDA is to provide an interactive or visual interpretation of the data that reveals knowledge that is otherwise indecipherable in the data's raw form (Hand, 2001). A good example of EDA is social network analysis where the output is a visual diagram with nodes representing individuals and links between nodes representing relationships. Social network diagrams help analysts decipher which groups of individuals make up the various cliques and who are the critical bridges between those cliques.

The second and third tasks are descriptive and prescriptive modeling, respectively. An example of descriptive modeling is cluster analysis, which can take a database of customer attributes and segment the customers into clusters of similarity usually for the purposes of direct marketing (Hand, 2001). Predictive modeling is similar to descriptive modeling except that there is a specific variable that is the target of the

study (Hand, 2001). In decision tree classification, which is an example of predictive modeling, one may be interested in deciphering which customer variables best predict the "satisfaction" variable value of "happy." This knowledge can guide managers in their decisions regarding which areas to improve if they want to better serve their customer base.

The fourth type of common data mining task is discovering patterns and rules (Hand, 2001). The classic example of this task is a form of association rule mining commonly referred to as market basket analysis (Han and Kambler, 2006). The retail industry uses market basket analysis on transaction data in order to infer which products sell together (in the same market basket). The results can be used to generate recommended buys for customers, guide catalogue design decisions, and describe customer shopping behavior (Han and Kambler, 2006). Market basket analysis is really just a fancy word that marketing consultants use for association rule mining, which is further explained in the next section of this chapter.

The last data mining task is one which most people use daily, whether they realize it or not: retrieval by context. This is the form of data mining employed by search engines like Yahoo® and Google™. This data mining task is also used by spam-filters to screen email and stop malicious web-page content.

Although the five tasks described above are clearly differentiated, it is easy to see that many of them share components or even whole functions (Hand, 2001). The five basic tasks are provided primarily to serve as a snap-shot of where data mining has been

applied. The most sophisticated data miners are customizing their algorithms to combine or even transcend the traditional categories.

Association Rules

Association rule mining was first motivated for the purpose of market basket analysis (Agrawal, et al., 1993). Thanks to the information age and the advent of the barcode, retailers have been collecting data on what is selling where and when for decades. In other words, businesses have been storing data about which products tend to be purchased together in the same "market basket." As the data storage and processing power of systems have been increasing exponentially, the amount of data available for market basket analysis has increased. So what exactly is an association rule? That question is best answered by presenting a famous example from market basket analysis: the beer and diapers rule (Berry and Linoff, 2004). The association rule describing the relationship between beer and diapers is represented as the Boolean vector

IF *buy beer*, **THEN** *buy diapers* (1)

where "beer" is the antecedent and "diapers" is the consequent. Similarly, an association rule can be represented in short hand as

$$Beer \Rightarrow Diapers\,. \tag{2}$$

Association Rule Metrics

There are three common metrics that accompany an association rule: support, confidence, and some measure of interestingness (Hand, 2001). These measures are critical in the evaluation phase of an association rule mining initiative. Support and confidence are measures of the strength of the given rule (Larose, 2005), whereas the "interestingness" factor is more like a measure of correlation (Han and Kambler, 2006).

Strength Measures

The support and confidence of a rule reflect the usefulness and certainty of the discovered rule, respectively (Han and Kambler, 2006). In equation form, the support of a rule is given by

$$\text{support}(A \Rightarrow B) = P(A \cap B) = \frac{\text{\# of transactions containing both A and B}}{\text{total \# of transactions}} \tag{3}$$

where *A* is the antecedent and *B* is the consequent (Larose, 2005). The confidence of a rule is expressed as

$$\text{confidence}(A \Rightarrow B) = P(B \mid A) \tag{4}$$

where *A* is the antecedent and *B* is the consequent (Larose, 2005). By the properties of probability, we also know that confidence can be expressed as

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{\# of transactions containing both A and B}}{\text{\# of transactions containing A}} \tag{5}$$

Returning to the beer and diapers example: Consider that there are 100 transactions in a grocery store where 20 patrons bought beer, 15 patrons bought diapers, and 10 of these transactions included both beer and diapers purchased together. Given the rule expressed in equation 1 ($Beer \Rightarrow Diapers$), the support is 10% and the confidence is 50%. This means that of the 100 transactions in our example, 10% involved the purchase of both beer and diapers, and 50% of the customers who bought beer, also bought diapers.

Equations 3, 4, and 5 not only represent the strength of the given rules, but they are also critical elements in the generation of the rules. For any given association rule mining application, the analyst will only consider rules that satisfy a minimum support and confidence threshold (Han and Kambler, 2006). These thresholds are used by the association rule mining algorithm to generate the results. There is no generic rule of thumb for determining the minimum support and minimum confidence levels for all association rule mining applications. The analyst must determine them based on the data and the desired outcome of the study. For example, an analyst performing market basket analysis for a grocery store may be satisfied with a support of 20% and a confidence of 70%. However, an analyst seeking fraudulent transactions may want to go after much lower levels of support, perhaps even lower than 1%, assuming that the overwhelming majority of transactions are not fraudulent (Larose, 2005).

Correlation Measures

Strong association rules are not necessarily interesting to users (Han and Kambler, 2006). Returning to the beer and diaper example, one can conclude that if a customer purchases beer there is a 50% probability (confidence) that the customer will also buy diapers. However, what if the probability of a customer buying diapers alone is 50% in the first place? Furthermore, what if the probability of a customer purchasing diapers alone is 60% or 70%? In that situation, adding the antecedent of beer actually reduces the chance that the customer will purchase diapers. For this reason, it is often necessary to produce some "interestingness" or correlation measure with each rule in order to rank the results (Han and Kambler, 2006).

The most common correlation measure used in classic market basket analysis is lift (Berry and Linoff, 2004). Lift is defined as the ratio of the rule's confidence to the expected confidence of the consequent (Larose, 2005). It can be expressed as

$$\text{lift}(A \Rightarrow B) = \frac{P(A \cup B)}{P(A)P(B)} = \frac{P(B \mid A)}{P(B)} = \frac{\text{confidence}(A \Rightarrow B)}{\text{expected confidence of } B} \tag{6}$$

where A is the antecedent, and B is the consequent. If the result of equation 6 is less than 1, then the occurrence of A is negatively correlated with B. If the result is equal to 1, then A and B are independent and there is no relationship between them. If the lift is greater than 1, then the occurrence of A is positively correlated with B and the rule is deemed "interesting" (Han and Kambler, 2006). Returning to the beer and diapers example, with

a confidence of 50% and an expected confidence in diaper purchase of 15%, the lift becomes

$$lift(Beer \Rightarrow Diapers) = \frac{P(Diapers \mid Beer)}{P(Diapers)} = \frac{0.5}{0.15} = 3.33\,. \tag{7}$$

Since the lift is greater than 1, one can conclude that the rule is interesting.

Another measure of correlation is importance, which is very similar to lift. Importance is used by Microsoft® SQL Server 2005 as a measure of "interestingness." The best way to describe importance is as the log of the ratio of the confidence of a rule to the expected confidence of the consequent given the absence of the antecedent (MSDN forum, 2006). Importance is expressed as

$$\text{importance}(A \Rightarrow B) = \log\left(\frac{P(B \mid A)}{P(B \mid \text{not } A)}\right). \tag{8}$$

Unlike lift, importance measures revolve around 0. Therefore, a result for equation 8 of less than zero means that A is negatively correlated to B, equal to zero means independence, and greater than zero means that A is positively correlated to B (MSDN forum, 2006).

Other correlation measures include χ^2, all-confidence, cosine (Han and Kambler, 2006), and J-measure (Hand, 2001; Larose, 2005). These measures were developed to compensate for the sensitivity of lift to massive amounts of null transactions in large databases (Han and Kambler, 2006). A null transaction is one that does not include the antecedent or the consequent. Because this study involves a relatively small dataset, these improved, but more complex, correlation measures are not necessary.

Association Rule Mining Case Studies

The most profound and effective uses of data mining are easily recognized. For example, Google™'s text-mining search engine is the enabler of their multi-billion dollar business model which has made them a giant among the high-tech firms of the 21st century. An example more closely related to this study might be Amazon or iTunes' use of market basket analysis. Today, most e-commerce consumers are accustomed to receiving emails or passing links to "recommended buys" that are in line with their purchasing preferences. This form of direct marketing is very effective and has transformed the business world (Economist, 2007). However, market basket analysis is not the only area where association rule mining applies.

Before attempting any data mining study, it is necessary to first review other studies to determine the acceptable format and get an idea of the appropriate application of the data mining technique. In this review of data mining case study literature, the focus is on applications that involve the use of association rules in areas other than classic market basket analysis. Unfortunately, no studies applying association rules to a points-based assessment system have been found in the existing literature. However, as demonstrated by the sample of case studies to follow, the wide breadth of areas and industries that data mining techniques have been applied to seems limitless, constrained only by the objectives of the researcher. This review is not intended to serve as a

comprehensive list of association rule mining case studies, but rather, a modest sample that sets the stage for its application to the LEED-NC building assessment system.

Personnel Selection

High-tech firms often find that it is difficult to acquire competent personnel to fill their needs. However, it can be even more difficult for these firms to retain their human capitol due to the changing nature of knowledge workers in the high-tech industry. Hsinchu Science Park, a semi-conductor foundry in Taiwan, turned to Chien and Chen (2008) to help them reengineer their recruiting process in order to attract a talented and more loyal employee base. The data set included employee demographics, performance, recruitment channel, and reason for leaving (if applicable). The researchers applied predictive data mining techniques to determine the employee attributes that were most commonly associated with high performance and low turnover. CHIAD, a classification algorithm that also produces association rules, was employed to build a decision tree from which association rules were derived. Several interesting association rules were discovered and incorporated in the recommendations for Hsinchu's new recruitment process. One of recurring associations was between an employee's recruitment source being external and short-term turnover. As a result, one of the researchers' recommendations was to increase internal recruitment initiatives (Chien and Chen, 2008).

Course Management Systems

With the increased use and proliferation of course managements systems like Blackboard, WebCT, and Moodle, a vast warehouse of data is generated that has the potential to give educators more immediate feedback into the use and success of online course content. Romero, Ventura, and Garcia (2007) provide an overview of data mining as it applies to e-learning. Their study provides examples of teaching applications from most of the available data mining techniques. For association rule mining, there were several very interesting applications suggested: from generating recommended learning activities and shortcuts to discovering relationships between student usage information and performance. The authors are very optimistic about the use of data mining in the e-learning environment and the ways that it could revolutionize education in the 21st century (Romero, et al., 2007).

Automotive Warranty

The best way to ensure profits from warranty sales is to minimize the costs and hassle of warranty-related repairs and replacements. Of course, the best way to minimize those liabilities is to build a product of high quality that is robust enough to last through the warranty period without failure. However, quality is a nebulous concept. A company could spend exorbitant amounts of money in pursuit of “quality” only to realize marginal improvement. So what processes or components of the product should the company

focus on to deliver quality improvements that will directly reduce warranty claims? This question was the focus of the data mining algorithm developed by Buddhakulsomsiri, Siradeghyan, Zakarian, and Li (2007) for automotive warranty data. The algorithm employed is custom-written for warranty data because it uses product features like engine type or production date as the predictor of the problem-related labor code of the warranty claim. This ensured that all the associations reported by the algorithm will result in a warranty claim. Although the algorithm was successful, the limited integration of data between the manufacturers, dealerships, and repair-houses resulted in association rules that provided limited new knowledge. No matter how powerful or robust the data mining algorithm is, it is no substitute for good data that is consolidated and considered complete (Buddhakulsomsiri, et al., 2007).

Library Circulation

Market basket analysis of retail transactions is the most common application of association rule mining. Library circulation can very easily be compared to retail transactions. In the same way that the retail industry has benefited from association rule mining, associations between books and media that are checked-out together may give library patrons more insight to guide their searches. It may also provide the library staff a tool for determining a better physical shelf layout. Apriori, the seminal algorithm used in association rule mining, was applied to a set of 20,000 checkout transactions from the library of the University of Waikato, New Zealand. The Waikato library used the Library

of Congress classification scheme to develop their shelf layout. With over 50,000 different titles in the data set, the researchers chose the two-letter classification sub-categories provided by the Library of Congress as their level of abstraction. This reduction of the scope resulted in a data set of 4,308 transactions that contained titles from more than one sub-category. The generated rules were evaluated and it was concluded that the existing shelf layout did not require the majority of the patrons to walk around excessively to find strongly associated titles. Additionally, recommendations were made to develop a system of signs on the shelves to point patrons in the general direction of associated works (Cunningham and Frank, 1999).

Aviation Maintenance

Mistakes in aviation maintenance have real life and death consequences. The aviation industry is ever-conscious of the perception of its safety record. For this reason, research into human factors in aviation maintenance work has become a popular topic. Zhang and Yang (2006) applied association rule mining to a set of data collected on 89 aviation workers with 891 records of mistakes. The data consisted of several employee characteristics including marital status, education, hobbies, age, and a subjective observation of their character. The Apriori algorithm was applied to reveal association rules, and only rules resulting in “low standard” as the attributed cause of the mistake were retained. The study produced a collection of employee attributes that serve as warning signs for managers to look out for. With this tool, the manager can

systematically evaluate the crew and intervene when he or she notices a preponderance of warning signs (Zhang and Yang, 2006).

Crime Data

The use of information systems to help police forces is not new, but integrating the vast array of government information systems that encompass the United States' police forces will help to advance their mission. Much research is being directed at developing digital government tools to approach that objective. One such tool is COPLINK, which is a digital government program developed at the University of Arizona to serve as a fully integrated crime information system for the state of Arizona. COPLINK has a wide array of built-in data mining capability. One of the data mining techniques applied is an association rule algorithm that automatically detects associations between crimes in the COPLINK system. Data stored by detectives at the scene of a crime are instantly checked for associations with other crimes of a similar type or similar circumstance. These automatic leads have a lot of potential and may reduce the time and burden placed on detectives to sort through thousands of dead ends before breaking a case (Chen, 2002).

III. Methodology

Association rule mining was used to determine the synergy between credits in the Leadership in Energy and Environmental Design for New Construction (LEED-NC) Building assessment system. Therefore, this section will begin with an outline of the basic inner workings of the seminal algorithm used in association rule mining: Apriori. Then there is a brief discussion of data mining software packages, including the one chosen for modeling associations between credits in LEED-NC. Finally, applying the Cross-Industry Standard Process for Data Mining (CRISP-DM) that was introduced in Chapter II, the iterative process used to develop the data mining model employed in this study will be presented.

The Apriori Algorithm

Association rules are simple enough to understand. The mathematics is relatively easy and involves applying rudimentary statistics. However, deriving strong and interesting rules from a large dataset is an entirely different matter. Even when limited to binary attributes and positive cases (e.g., *buy diapers = yes*), the number of possible association rules is

$$\#\text{ of Rules} = k \times 2^{k-1} \tag{9}$$

where k is the number of items (Larose, 2005). Suppose, for example, that we have a tiny grocery store that only sells 100 products. According to Equation 9, there are 6.4×10^{31} association rules to generate. Even with modern computer processing power, that is a

daunting task. One can only imagine generating all the possible association rules for a chain retailer that sells tens of thousands of products and processes hundreds of thousands of transactions a day. To manage the computations, Agrawal and Srikant (1994) introduced Apriori, which became the seminal algorithm for mining Boolean association rules (Han and Kambler, 2006; Agrawal, et al., 1994).

In general terms, association rule mining can be broken down into two steps (Han and Kambler, 2006):

1. **Find all frequent itemsets:** Frequent itemsets are defined as any combination of items that occur at least as frequently as the user-defined minimum support threshold.
2. **Generate strong association rules from the frequent itemsets:** The generated rules must satisfy the user-defined minimum confidence and minimum support thresholds.

The rest of this section describes how the Apriori algorithm accomplishes these steps.

Discovering Frequent Itemsets (Step 1)

The key to the Apriori algorithm is its use of the Apriori property. The Apriori property states that if an itemset is not frequent, then the combination of that itemset and any other item or itemset is also not frequent (Larose, 2005). This means that the algorithm does not have to calculate the frequency of any itemset that has a non-frequent proper sub-set (Hand, 2001).

Consider L_k, a vector containing all the itemsets made up of k-items that occur often enough to satisfy the minimum support threshold. The algorithm begins by forming L_1, then the algorithm uses L_1 to generate L_2 and so on until all the frequent itemsets are found. This is accomplished by first generating C_k, a vector of candidate itemsets, by joining L_{k-1} with itself. Then, C_k is pruned using the Apriori property and the surviving itemsets become L_k (Larose, 2005). This is represented in algorithm form in Figure 3.

```
i=0;
Ci = {{A} | A is a variable};
while Ci is not empty do
        database pass:
                for each set in Ci, test whether it is frequent;
                let Li be the collection of frequent sets from Ci;
        candidate formation:
                let Ci+1 be those sets of size i+1 whose all subsets are frequent;
end.
```

Figure 3. Algorithm for Frequent Itemset Discovery in Apriori algorithm (Hand, 2001)

Generating Association Rules (Step 2)

Once the frequent itemsets are discovered, generating strong association rules from them is a straightforward process (Han and Kambler, 2006). First, for each frequent itemset, l, the algorithm generates all nonempty subsets s of l. Then, for each subset, s, the algorithm will output the rule "$s \Rightarrow (l - s)$" as long as the following rule is satisfied:

$$\text{confidence}\big(s \Rightarrow (l - s)\big) = \frac{\#\,\text{of transactions containing } l}{\#\,\text{of transactions containing } s} \geq Min\_Con\,, \qquad (10)$$

where *Min_Con* is the minimum confidence threshold (Han and Kambler, 2006). The left-hand side of equation 10 is a modified version of confidence which was introduced in Equations 4 and 5. Because the rules are formed from the frequent itemsets discovered in step one, they already satisfy the minimum support threshold.

Alternative Methods of Rule Generation

There are many association rule mining algorithms that are newer than the Apriori algorithm (Han and Kambler, 2006). These updates were written primarily to improve efficiency for the purposes of scaling the algorithm to tackle extremely large databases (Hand, 2001). Since the dataset for this study is relatively small, the Apriori algorithm is the best because, although it is inefficient, it is the most robust. It is also possible to derive rules from classification mining results such as decision trees (Hand, 2001). However, these methods are based in predictive modeling which is outside of the scope of this study.

Software Selection

There are numerous data mining software packages to choose from. Full-service packages like SPSS Inc.'s Clementine (Clementine, 2008) and IBM's DB2 Intelligent Miner (Miner, 2008) are expensive, but offer the full range of data mining capabilities mentioned in Chapter II. They are also very powerful and backed up by the support and development of top software engineering firms, which is what makes them so expensive.

These name-brand data mining packages are typically used by business intelligence consultants and professional data mining practitioners. There are also free full-service packages developed by academics, like Weka (Weka, 2008) and Keel (Keel, 2008). They are less user-friendly and do not have all the capability and support as the name-brand packages. There are also various stand-alone association rule mining programs available on independent data mining websites like KDNuggets (KDNuggets, 2008). These programs are typically available for download as source java code. Once the code is downloaded, most of the stand-alone association-rule mining programs have a user-friendly graphical-user interface. It is important to understand the intent of the programmer that posted the algorithm. Often, the program is posted to display its efficiency in comparison to other programmers' offerings. If that is the case, there may have been short-cuts taken in the code in order to optimize speed at the cost of performance and robustness. The type of data mining package chosen for this study is an add-on to a database management system (DMS). Both Oracle (Oracle, 2008) and Microsoft (MS) SQL Server 2005 (MS SQL, 2008) offer data mining capability integrated with their DMS. Microsoft offers a simplified version of their add-on for free, which is the primary reason it was chosen for this study. The Microsoft data miner works as an add-on in MS Excel that calls on functionality from MS SQL Server 2005 Analysis Services. Because the free version offers an association rule miner that uses the Apriori algorithm, it is the optimal tool for this study.

LEED Data Mining Process

This section describes the iterative process used to develop the model used to mine association rules in the LEED-NC version 2.0 and 2.1 building assessment systems. Following the CRISP-DM format, this section starts with a brief discussion of the business and data understanding phases, followed by data preparation, modeling, evaluation, and deployment. It is important to note that this process is iterative; although the phases are presented in sequential order, they were not necessarily performed in that order. In fact, several of the phases were revisited and constraints added as the model developed. For more information on the nature of a typical CRISP-DM reference model please refer to Figures 1 and 2 in Chapter II.

Business and Data Understanding

The majority of the business understanding phase of this study was accomplished in the preparation of the background and literature review for Chapters I and II of. The business understanding led to the development of the research questions that were presented in Chapter I are reproduced below.

1. What are the most interesting credit bundles?
2. Are there any actionable implications of the revealed credit synergies?
3. Can the USGBC use association rules to streamline the credit review process?

Please refer to Chapter I for more description on the development of these research questions.

The data provided for the study was assembled by a member of the USGBC's Technical Development Team on November 7th, 2007. Their team uses a series of Excel spreadsheets to track aggregate trends in credit accomplishment. The data is presented in a tabular format with each row representing a LEED-NC construction project. Table 4 provides an excerpt from the USGBC's LEED credit achievement tracking spreadsheet. In the interest of protecting the USGBC's proprietary rights, a full version of the raw data is not available. Each column in the table represents 1 of the 69 LEED points. For each project, the cell corresponding to each credit is filled with a 1 if the project attained that credit, a -1 if the project attempted the credit but was denied, and left blank if the project did not attempt the credit.

Table 4. Excerpt from LEED credit achievement tracking data.

	SUSTAINABLE SITES													
ID #	**1**	**2**	**3**	**4.1**	**4.2**	**4.3**	**4.4**	**5.1**	**5.2**	**6.1**	**6.2**	**7.1**	**7.2**	**…**
73	1			1	1		1	1	1		1	1		…
94	1	1	1	1			-1	1	1	1		1	1	…
113	1	-1	1	1	1	-1	1	1	1	1	1	1	-1	…
139	1			1	1		1	1	1	1	1	1		…
195	1			1	-1		1	-1	1	1	1		1	…
203	1				1		-1		1	-1			1	…
209	1							1	1	1	1	1		…
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

The data provided was from all construction projects that have been certified under LEED-NC versions 2.0 and 2.1. Projects from Version 1.0 are omitted because it was a pilot program that does not resemble the current 69-point LEED system.

Additionally, there are so few certified projects from LEED-NC version 2.2 that it is omitted to preserve the integrity of the study. There were several changes to the wording of certain credits between versions 2.1 and 2.2. The change of language my induce credit synergies specific to version 2.2 and eliminate synergies discovered in 2.0 and 2.1. Therefore, the reasons for including both 2.0 and 2.1 in the study are two-fold: (1) The most significant change in the update to version 2.1 was the introduction of LEED Online for project reporting; the wording of the credits themselves was changed very little. (2) The inclusion of both versions provides 764 projects for the model. Omission of either version 2.0 or 2.1 would effectively halve the population which would reduce the support metrics for all of the rules being mined.

Data Preparation

One of the most important decisions in an association rule mining model is what level of concept abstraction to choose for analysis (Cunningham and Frank, 1999). Association rules between all 69 points in LEED-NC could have been mined, but such an analysis would produce a lot of trivial results. In order to limit the model to interesting associations between differentiable credits, the prerequisites and Innovation and Design (ID) credits are omitted. This was done because all projects achieve the prerequisites and it is nearly impossible to differentiate the cause of synergies between the miscellaneous innovation and design credits.

Recall from the discussion in Chapter II in the section titled "Structure of LEED" that there is a difference between points and credits. Points are the actual scores that are tallied for the attainment of a certain level of certification. Some credits are worth only one point, whereas others are worth several points. To eliminate redundancy, all points that represent increments of the same design implementation are omitted, leaving only the base point of that credit for inclusion in the model. This is to avoid trivial association rules between two points that are merely increments of the same design implementation. It is important to note that points within the same credit that represent unique green amenities and not merely increments of the same design implementation have been included in the model. Refer to Appendix B for a complete display of included and omitted credits.

Energy and Atmosphere Credit 1 (EA-1) requires special attention with regards to inclusion in the association rule mining model. Although it is only one credit, it is worth a possible 10 points, making it the most significant credit in the LEED-NC building assessment system. It cannot be treated like the other multi-point credits because there is a very big difference between a construction project that achieves 1 point under EA-1 and a building that earns 10 points. High-scoring projects in EA-1 are facilities with cutting-edge energy systems and building envelope design. Low-scoring projects in EA-1 are facilities that avoid energy efficiency in favor of other environmental advocacy issues. To compensate for the difference, EA-1 is split into three collective-point credits. "EA-1 Low" is the "credit" recognized by the model for every project that achieves 1 to 3 points in EA-1, "EA-1 Middle" is the label for projects that earn 4 or 5 points, and "EA-1 High"

is granted to projects achieving 6 to 10 points. The reason for the selected point distribution is that each group represents approximately one-third of the total number of projects in the dataset that achieved points in EA-1. Refer to Appendix B for a complete display of the included, omitted, and collated credits.

The Microsoft data mining software used to build the model accepted the data in the tabular form represented in Table 4. However, including the "-1" entries for failed credits is outside the scope of this research. These entries are converted to blanks for the model. The resulting input is a table with the original 764 rows of projects and 47 columns representing the unique credits listed in Appendix B.

Modeling

Before the Apriori algorithm will mine association rules from the input data, the user must first define the minimum support and confidence thresholds. Because there are no historic applications of association rule mining to LEED-NC data or anything similar, the first iteration was performed with a minimum support of 20 items and a minimum confidence of 50%. In other words, the algorithm generates all possible itemsets that have at least 20 instances out of the possible 764 projects. In the dataset, the least commonly achieved credit was Materials and Resources Credit 1.3 (MR-1.3) which was only earned by 21 projects. Setting a minimum support of 20 occurrences ensures that MR-1.3 and no other credits are systematically left out of the analysis. A minimum confidence of 50% ensures that the algorithm reports all associations where the majority

of the instances of the antecedent also include the consequent. The concept is to start with very wide criteria and to narrow them to the most interesting association rules.

Using these criteria for the first iteration, there were so many results (i.e., association rules) that the algorithm could not produce them all. The association rules that were produced were very large, yielding rules with 4 and 5 antecedents. This is not surprising considering the fact that there are only 47 possible items with which to make an itemset and that a project cannot be certified unless it scores at least 26 points with the majority of projects scoring higher than that. However, large association rules are redundant. They produce entire series of rules where the same antecedents are present, just in different positions in the rule. There is also a very limited amount of knowledge that can be gained from large rules because most design professionals do not think 5 credits at a time. This assertion is reflected in the GSA's cost study which identified credit synergies in groups of 2 or 3. Fortunately, the Microsoft data mining software used to build the model also has a maximum itemset size threshold. Therefore, the second iteration is the same as the first but with a maximum itemset size of 3 items. This limits the results to association rules with 1 or 2 antecedents and 1 consequent.

The second iteration produced 18,490 rules with importance values as high as 0.99. Because the importance value is a measure of correlation, it is the metric of interest in determining whether or not the credits are synergistic. Subsequent iterations with higher minimum support and minimum confidence thresholds were attempted; however, they yielded low importance values close to zero which provides weak evidence of a true relationship. In order to maximize the importance values, the second iteration of the

model is selected to serve as the LEED-NC association rule mining model for analysis. To reduce the number of results to a manageable group for evaluation, the results were sorted in terms of importance from highest to lowest. Figure 4 shows a graph of the importance values for the top 3000 rules generated by the model.

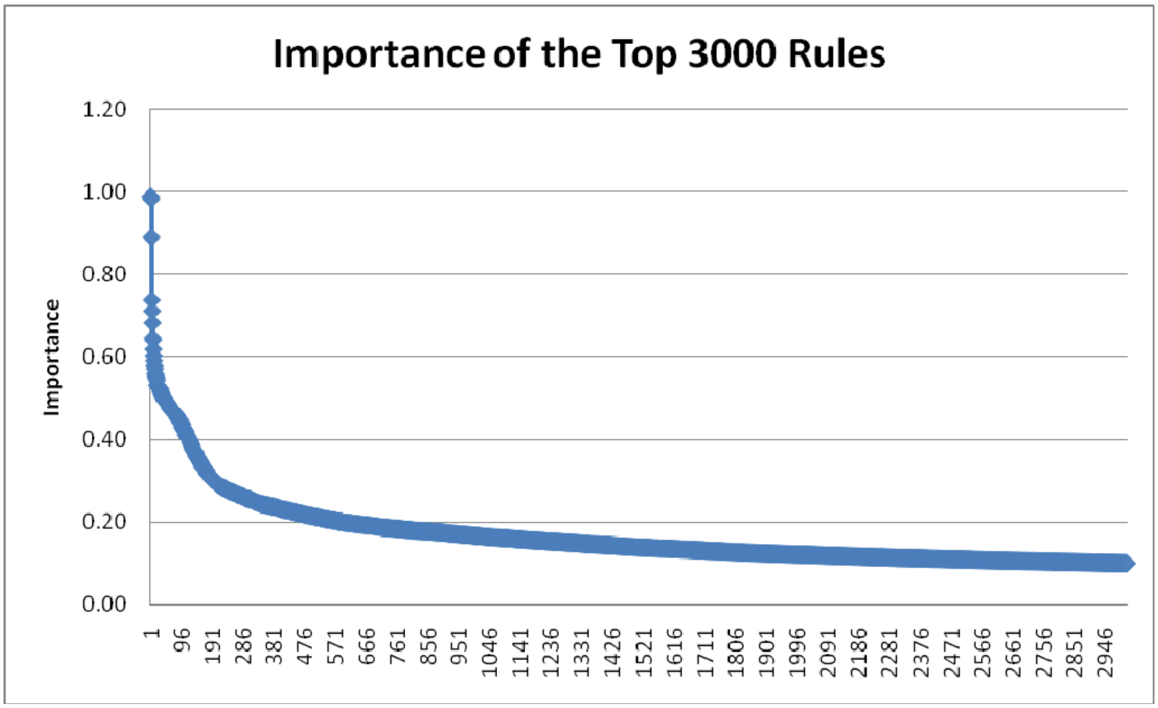

Figure 4. The 3000 highest importance values in the association rule model

There were 641 association rules with importance greater than 0.20. This cutoff value was arbitrary, but it provided a manageable number of rules with which to begin the evaluation phase. The iterative process used to select the minimum support and minimum confidence thresholds and the constant refining of the model blurs the line

between the modeling and evaluation phase of a data mining process. Several failed attmepts and incorrect tangents were left-out of this discussion, but are an unavoidable part of the CRISP-DM approach.

Evaluation

The 641 association rules produced by the model were still too vast to extract any usable knowledge from them. Therefore, the first step in evaluating them is to reduce the results down to the strongest and most interesting rules. Barry and Linoff (2004) provide a basis of classifying association rule mining results as actionable, trivial, or inexplicable. They mention that the majority of association rule mining results are either trivial or inexplicable, meaning that they either reveal knowledge of little value or appear to be anomalous. Only a very few results will have actual and actionable implications. However, their categories do not account for rules that are redundant or of unknown value.

The results from the LEED-NC credit association model were categorized using a format similar to Berry and Linoff (2004). Rules were classified under one of the following labels: interesting, repeat, trivial, or inexplicable. Rules that combined credits that the researcher does not think could possibly come from the same sustainable design strategy were labeled "inexplicable." Any rule that combined points within the same credit were classified "trivial". Whenever a rule was encountered that was the exact reverse of a previously cited rule (e.g., $X \rightarrow Y$ and $Y \rightarrow X$), it was labeled "repeat."

Taking a very liberal definition of the word, everything that did not fall under the other three categories was considered “interesting.” This method of classifying the results provided 135 “interesting” association rules.

To reduce the remaining 135 results, the 3-item association rules were grouped into recurring synergies. A recurring synergy is a 2–item association that recurs often in many other 3-item rules. It is reasonable to assume that the strength of the 2-item association was dominating the importance of the rule and causing it to appear with other credits in the antecedent. Once the groups were formed, one rule from the recurring synergy was selected to represent the group. This process reduced the results to 56 interesting and recurring credit synergies. A table of all 641 results with their categorical labels and recurring synergy groups can be found in Appendix C.

The final step in the evaluation phase was to select a group of the most interesting association rules to present to the USGBC in order to get their input on the reason for the credit synergy and any actionable implications they believe it may have revealed. From the 56 candidate rules, 24 were selected and presented to the LEED Technical Development Team. A transcript of their input can be found in Appendix D. This part of the results evaluation was the most subjective. However, it was necessary in order to reduce the number of credit bundles to a group that could be analyzed meaningfully in the time available with the USGBC staff.

Deployment

This thesis document represents the “deployment” deliverable of the LEED-NC association rule mining process. It is left to the USGBC to determine how to employ the knowledge gained by this data mining exercise. A full discussion of conclusions and future research areas is provided in Chapter V.

IV. Results

This chapter reports the results of the Leadership in Energy and Environmental Design for New Construction (LEED-NC) association rule mining model. Discussion based on input from the LEED Technical Development Team on the reasons and implications of the results is provided. Finally, a short discussion on the applicability of these rules to streamlining the United States Green Building Council's (USGBC) credit review process will be presented.

Interesting Rules

This section addresses research question 1: What are the most interesting credit bundles? An important definition in this question is what exactly is meant by the word "interesting." An association rule was deemed interesting if the credits appear to contribute to the same sustainable design strategy. This concept is best demonstrated by the "trivial" rules revealed by the model. Table 5 presents association rules that include LEED points from the same credit. It can be reasonably assumed that these credit bundles are achieved as part of the same sustainable design strategy by the fact that they address the same sustainability issue. Therefore, the existence of these "trivial" rules among the results of the model lends credibility to the assertion that association rules are the appropriate indicator of credit synergy. However, they do not constitute new

knowledge to either green design practitioners or the USGBC, hence they are labeled trivial.

Table 5. Trivial Association Rules

Confidence	Importance	Rule
100%	0.99	MR 1.3 -> MR 1.1
67%	0.89	EQ 7.1 -> EQ 7.2
73%	0.56	EQ 6.2 -> EQ 6.1
75%	0.50	EQ 3.1 -> EQ 3.2
59%	0.30	SS 6.1 -> SS 6.2
83%	0.23	SS 5.1 -> SS 5.2
51%	0.21	EQ 4.2 -> EQ 4.4

The confidence values are listed in Table 5 for reference, but recall that the importance factor is the critical metric in determining the synergy of the credits because it is a measure of correlation. The most interesting rules resulting from the association rule mining analysis will demonstrate sustainable design strategies that bring together points from different credits and credit categories. Therefore, the 24 most interesting rules were selected from the results to be reviewed by the LEED Technical Development Team. They are listed in Table 6 and ranked by importance factor.

Table 6. Interesting Rules

#	Confidence	Importance	Rule
1	74%	0.50	EA 2.1 -> EA 1 High
2	55%	0.37	EA 2.1, SS 6.1 -> SS 5.1
3	55%	0.37	EA 2.1, SS 6.2 -> SS 5.1
4	50%	0.36	EQ 6.2, EQ 3.2 -> EQ 2
5	61%	0.27	EQ 1 -> EQ 7.2
6	53%	0.27	EQ 6.2, SS 8 -> EA 1 High
7	52%	0.26	EQ 2, EQ 6.1 -> EA 1 High
8	50%	0.25	EA 2.1, EA 1 High -> EQ 6.1
9	74%	0.24	MR 6, SS 7.1 -> SS 6.1
10	71%	0.24	WE 2, SS 5.1 -> SS 6.2
11	83%	0.24	MR 6, EQ 3.2 -> EQ 4.4
12	63%	0.23	SS 5.1, WE 1.1 -> SS 6.1
13	69%	0.23	EA 2.1, EA 1 High -> EQ 8.1
14	95%	0.23	SS 2 -> SS 4.1
15	64%	0.22	EA 2.1, SS 4.3 -> EA 6
16	90%	0.22	WE 2, EA 5 -> EA 3
17	70%	0.22	SS 5.1, MR 7 -> EQ 8.1
18	77%	0.21	EQ 2, EA 5 -> EQ 7.2
19	70%	0.21	MR 7, EQ 1 -> EQ 4.4
20	62%	0.21	WE 2 -> SS 6.1
21	64%	0.21	WE 2, SS 5.2 -> SS 6.2
22	68%	0.20	MR 6, SS 7.2 -> SS 6.1
23	61%	0.20	MR 4.1 -> SS 7.2
24	51%	0.20	SS 5.2 -> SS 6.1

The USGBC's Technical Development Team agreed with the high level of interest in all the rules in Table 6 except rules 9, 22, and 23, which they deemed inexplicable. The reasons they generated for the appearance of the rules spanned a wide range from energy-focused project priorities to the probable location of the project. A transcript of their full input is available in Appendix D of this report.

Discussion of Implications

This section addresses research question 2: Are there any actionable implications of the revealed credit synergies? The discussion generated by the presentation of the results from Table 6 did reveal some very general implications about LEED. It verified some intended interaction between certain credits. Responses of "expected" and "not surprising" to rules 1, 5, 16, and 18 reinforced that these credit synergies were intentionally built into the system. Additionally, there were several inferences to credit synergies which revealed that an "experienced" and "integrated design team" was clearly necessary to accomplish the credit synergy. This response can be found to rules 4, 6, 8, and 18. Both rules involving Indoor Environmental Quality Credit 7.2 (EQ 7.2) helped to reinforce decisions made when LEED was updated to version 2.2. Association rules 5 and 18 include credit synergies that may not exist in version 2.2 because the requirement to include monitoring systems for achievement of EQ 7.2 was removed to eliminate redundancy.

Perhaps the most interesting implication of these association rule mining results was the fact that the credit synergies appear to transcend categorical boundaries. Only 5

of the 24 rules listed in Table 6 are credit bundles of the same core category. A broader view of the full results in Appendix C also supports this assertion. The heterogeneous nature of the credit bundles supports the USGBC's vision for LEED as an evaluation of environmental performance from the "whole building perspective" (USGBC, 2006).

LEED Credit Review Process

This section addresses research question 3: Can the USGBC use association rules to streamline the credit review process? One of the motivations behind this research was the potential to develop credit bundles of such high confidence that the USGBC could forego the review of certain credits once its "sister" credit was approved. This application would expedite the credit review process. Unfortunately, the credit bundles revealed by the LEED-NC association rule mining model cannot be used in this way.

The confidence values for the association rules are not strong enough to support the assertion that the credit bundles are systematic. Only 1 of the 24 rules in Table 6 has confidence higher than 90%. Looking at the results more broadly, only 55 of the 641 rules in Appendix C have confidence higher than 90%. Recall that in the evaluation phase of the data mining process, the results were sorted in terms of importance and eliminated all rules with importance values less than 0.20. This act removed a lot of high-confidence rules. However, because those rules had such low values for importance it can be assumed that their existence was more a function of the popularity of the individual credits and not necessarily the strength of the relationship between the credits. Furthermore, the results from the LEED-NC association rule mining model were

classified as “interesting” according to a very subjective and human-error prone method. Even with the USGBC’s hypothesized reasons for the rules, there was simply not enough empirical evidence to support the use of these rules in systematically approving credits without review.

V. Conclusions

This chapter presents the conclusions reached during the research. It begins with general conclusions about the research and its application. Then, recommendations for future research are presented. Finally, there is a brief section on the impact of the Leadership in Energy and Environmental Design for New Construction (LEED-NC) association rule mining model.

Conclusions of Research

This research successfully demonstrated the application of a data mining technique to a green building assessment system. More specifically, it demonstrated the use of an association rule mining algorithm on LEED-NC version 2.0 and 2.1 systems. The results showed that credit synergy clearly exists and that some credit bundles do emerge from the dataset. This suggests that there is a preponderance of sustainable design practitioners that are choosing credits based on whole-building concepts and not in discipline or environmental impact vacuums.

One tempting application of the revealed credit synergies is for design professionals to bring them into the schematic design-phase discussions about credit selection; however, this would be a serious misuse of the research. While the revealed credit bundles are interesting and spark discussion about credit interplay and its role in high-scoring sustainable design strategies, they are not intended to interfere with the established process of project credit selection. Every construction project has unique

needs based on several variables from geographic location to owner preferences. It would be a mistake to forego any of these needs purely for the sake of seeking “proven” credit synergies.

As with much research, the credit bundles revealed in this research represent only the credit synergies that are present in the data. However, the subjective methods employed in reducing the results to “interesting” credit bundles potentially eliminated several credit synergies that are present in the data. This study does not identify credit synergies that are unpopular or whose existence is obscure, but that does not mean that they do not exist. It was never the intention of this research to provide a comprehensive list of all possible credit synergies.

Recommendations for Future Research

Often, the most important output of a data mining model is the generation of more intelligent and focused research questions. This is integral to the Cross-Industry Standard Process for Data Mining (CRISP-DM) reference model first displayed in Figure 1 of Chapter II. In the reference model, there is an arrow that links the evaluation phase directly back to the first phase of the process (business understanding). The experience gained during this data mining process demonstrated the very real existence of this feedback.

Predictive Model for Streamlining the Credit Review Process

The United States Green Building Council (USGBC) needs a tool that can help streamline their credit review process; however, associations between credits are not sufficient for this application. The USGBC needs a predictive data mining model that is customized to LEED credit achievement data. This model should incorporate historic credit achievement, credit failure, and other variables such as geographic location and building type. Also critical in the development of this model is that the dataset used to generate it be from the same version of LEED as the project it is attempting to predict. A predictive model built on data from version 2.2 projects should not be used to predict credit achievement for projects in the newly proposed version 3.0 system. If there is sufficient data to build such a model, it could be incorporated into a Bayes network with a user-friendly graphical display. A Bayes network will represent each credit as a node in a network and each node could be activated if the credit is approved and deactivated if not. This tool could be used to predict what credits are expected to be achieved based on project variables and credits that have already been reviewed and approved. Based on the probability of those predictions, credit reviewers may decide to forego the review of expected credits.

Credit Review Process Improvement

Although the previous recommendation would provide a wonderful tool for the streamlining of the LEED credit review process, it does not take into account the entire

process. Any adjustments made to a business process should be analyzed from a systems perspective that will provide a holistic view. There are many business process improvement methodologies extant in the management literature. Selecting one and applying it in a way that is appropriate for the LEED credit review process is a critical step in developing the optimal solution for the USGBC.

Impact

Although this research does not provide tools for existing sustainable design practice or USGBC business processes, it does provide an interesting perspective on sustainable design strategy. Analyzing credit achievement at the aggregate level and then hypothesizing the design strategy based on revealed credit associations is unprecedented. It helps reinforce some of the basic underpinnings of whole-building design and shows that there are a plethora of strategies possible in the attainment of LEED-NC certification.

US Air Force

The Air Force conducts its Military Construction (MILCON) program under a heavily constrained budget that must compete with higher operational priorities and sister service needs. Now that the Air Force is enforcing its mandate on the use of LEED design principles on all its MILCON projects, there is a perceived demand for a simplified approach to LEED project delivery. This research builds on the list of common credit synergies provided by the General Services Administration (GSA).

However, it also reveals that there is no such thing as a common LEED design strategy that is systematically employed by the sustainable development community. There are no short-cuts in LEED project delivery. Therefore, the Air Force will have to continue to rely on the best-practices established by experienced sustainable design professionals.

USGBC

Although this research did not provide a tool for streamlining the USGBC's credit review process, it does offer some interesting insight into the development of LEED. It helped to reinforce some decisions that were made in the change from version 2.1 to version 2.2. Most importantly, it raises questions about credit synergies that LEED's Technical Development team expected to see in the results, but did not. This research's greatest impact on the USGBC is that it lays the foundation for the application of data mining techniques to future LEED data sets.

Sustainable Development Community

The impact this research offers the Air Force also applies to the greater sustainable development community. That is, that no one strategy is systematically employed by sustainable design professionals in the pursuit of LEED certification. Furthermore, the heterogeneous nature of the results reinforces the assertion that an integrated design team is necessary for successful LEED project delivery. Ultimately, the

most important impact of the identified credit bundles is that they support the vision of LEED as a system that rewards the fundamental concept of whole-building design.

Appendix A: LEED-NC v2.2 Building Assessment System

All 69 points from the LEED-NC Version 2.2 Building Assessment System are provided in the following checklist which is taken from USGBC (2008):

LEED for New Construction v2.2
Registered Project Checklist

Project Name:
Project Address:

Yes	?	No			Sustainable Sites	14 Points
Y			Prereq 1	**Construction Activity Pollution Prevention**		Required
			Credit 1	**Site Selection**		1
			Credit 2	**Development Density & Community Connectivity**		1
			Credit 3	**Brownfield Redevelopment**		1
			Credit 4.1	**Alternative Transportation**, Public Transportation Access		1
			Credit 4.2	**Alternative Transportation**, Bicycle Storage & Changing Rooms		1
			Credit 4.3	**Alternative Transportation**, Low-Emitting & Fuel-Efficient Vehicles		1
			Credit 4.4	**Alternative Transportation**, Parking Capacity		1
			Credit 5.1	**Site Development,** Protect or Restore Habitat		1
			Credit 5.2	**Site Development,** Maximize Open Space		1
			Credit 6.1	**Stormwater Design,** Quantity Control		1
			Credit 6.2	**Stormwater Design,** Quality Control		1
			Credit 7.1	**Heat Island Effect,** Non-Roof		1
			Credit 7.2	**Heat Island Effect,** Roof		1
			Credit 8	**Light Pollution Reduction**		1

Yes	?	No			Water Efficiency	5 Points
			Credit 1.1	**Water Efficient Landscaping**, Reduce by 50%		1
			Credit 1.2	**Water Efficient Landscaping**, No Potable Use or No Irrigation		1
			Credit 2	**Innovative Wastewater Technologies**		1
			Credit 3.1	**Water Use Reduction**, 20% Reduction		1
			Credit 3.2	**Water Use Reduction**, 30% Reduction		1

			Energy & Atmosphere		17 Points
Y			Prereq 1	**Fundamental Commissioning of the Building Energy Systems**	Required
Y			Prereq 2	**Minimum Energy Performance**	Required
Y			Prereq 3	**Fundamental Refrigerant Management**	Required

***Note for EAc1**: All LEED for New Construction projects registered after June 26th, 2007 are required to achieve at least two (2) points under EAc1.

			Credit 1	**Optimize Energy Performance**	1 to 10
				10.5% New Buildings or 3.5% Existing Building Renovations	1
				14% New Buildings or 7% Existing Building Renovations	2
				17.5% New Buildings or 10.5% Existing Building Renovations	3
				21% New Buildings or 14% Existing Building Renovations	4
				24.5% New Buildings or 17.5% Existing Building Renovations	5
				28% New Buildings or 21% Existing Building Renovations	6
				31.5% New Buildings or 24.5% Existing Building Renovations	7
				35% New Buildings or 28% Existing Building Renovations	8
				38.5% New Buildings or 31.5% Existing Building Renovations	9
				42% New Buildings or 35% Existing Building Renovations	10
			Credit 2	**On-Site Renewable Energy**	1 to 3
				2.5% Renewable Energy	1
				7.5% Renewable Energy	2
				12.5% Renewable Energy	3
			Credit 3	**Enhanced Commissioning**	1
			Credit 4	**Enhanced Refrigerant Management**	1
			Credit 5	**Measurement & Verification**	1
			Credit 6	**Green Power**	1

continued…

Yes	?	No	Materials & Resources		13 Points
Y			Prereq 1	**Storage & Collection of Recyclables**	Required
			Credit 1.1	**Building Reuse**, Maintain 75% of Existing Walls, Floors & Roof	1
			Credit 1.2	**Building Reuse**, Maintain 100% of Existing Walls, Floors & Roof	1
			Credit 1.3	**Building Reuse**, Maintain 50% of Interior Non-Structural Elements	1
			Credit 2.1	**Construction Waste Management**, Divert 50% from Disposal	1
			Credit 2.2	**Construction Waste Management**, Divert 75% from Disposal	1
			Credit 3.1	**Materials Reuse**, 5%	1
			Credit 3.2	**Materials Reuse**,10%	1
			Credit 4.1	**Recycled Content**, 10% (post-consumer + ½ pre-consumer)	1
			Credit 4.2	**Recycled Content**, 20% (post-consumer + ½ pre-consumer)	1
			Credit 5.1	**Regional Materials**, 10% Extracted, Processed & Manufactured Regionally	1
			Credit 5.2	**Regional Materials**, 20% Extracted, Processed & Manufactured Regionally	1
			Credit 6	**Rapidly Renewable Materials**	1
			Credit 7	**Certified Wood**	1
Yes	?	No			

Yes	?	No	Indoor Environmental Quality		15 Points
Y			Prereq 1	**Minimum IAQ Performance**	Required
Y			Prereq 2	**Environmental Tobacco Smoke** (ETS) **Control**	Required
			Credit 1	**Outdoor Air Delivery Monitoring**	1
			Credit 2	**Increased Ventilation**	1
			Credit 3.1	**Construction IAQ Management Plan**, During Construction	1
			Credit 3.2	**Construction IAQ Management Plan**, Before Occupancy	1
			Credit 4.1	**Low-Emitting Materials**, Adhesives & Sealants	1
			Credit 4.2	**Low-Emitting Materials**, Paints & Coatings	1
			Credit 4.3	**Low-Emitting Materials**, Carpet Systems	1
			Credit 4.4	**Low-Emitting Materials**, Composite Wood & Agrifiber Products	1
			Credit 5	**Indoor Chemical & Pollutant Source Control**	1
			Credit 6.1	**Controllability of Systems**, Lighting	1
			Credit 6.2	**Controllability of Systems**, Thermal Comfort	1
			Credit 7.1	**Thermal Comfort**, Design	1
			Credit 7.2	**Thermal Comfort**, Verification	1
			Credit 8.1	**Daylight & Views**, Daylight 75% of Spaces	1
			Credit 8.2	**Daylight & Views**, Views for 90% of Spaces	1

Yes	?	No	Innovation & Design Process		5 Points
			Credit 1.1	**Innovation in Design**: Provide Specific Title	1
			Credit 1.2	**Innovation in Design**: Provide Specific Title	1
			Credit 1.3	**Innovation in Design**: Provide Specific Title	1
			Credit 1.4	**Innovation in Design**: Provide Specific Title	1
			Credit 2	**LEED® Accredited Professional**	1

Yes	?	No	Project Totals (pre-certification estimates)	69 Points

Certified: 26-32 points, **Silver:** 33-38 points, **Gold:** 39-51 points, **Platinum:** 52-69 points

Appendix B: Credits Included in the Model

Table B-1 displays the credits used by the model to generate association rules. Points that were omitted are crossed-out with a line and points that were combined into a special collection of points are drawn in a box. (Table begins on next page)

Table B-1. Credits Included in Association Rule Mining Model

SUSTAINABLE SITES		
Credit 1	Site Selection	
Credit 2	Development Density & Community Connectivity	
Credit 3	Brownfield Redevelopment	
Credit 4.1	Alternative Transportation, Public Transportation Access	
Credit 4.2	Alternative Transportation, Bicycle Storage & Changing Rooms	
Credit 4.3	Alternative Transportation, Low-Emitting & Fuel-Efficient Vehicles	
Credit 4.4	Alternative Transportation, Parking Capacity	
Credit 5.1	Site Development, Protect or Restore Habitat	
Credit 5.2	Site Development, Maximize Open Space	
Credit 6.1	Stormwater Design, Quantity Control	
Credit 6.2	Stormwater Design, Quality Control	
Credit 7.1	Heat Island Effect, Non-Roof	
Credit 7.2	Heat Island Effect, Roof	
Credit 8	Light Pollution Reduction	
WATER EFFICIENCY		
Credit 1.1	Water Efficient Landscaping, Reduce by 50%	
~~Credit 1.2~~	~~Water Efficient Landscaping, No Potable Use or No Irrigation~~	
Credit 2	Innovative Wastewater Technologies	
Credit 3.1	Water Use Reduction, 20% Reduction	
~~Credit 3.2~~	~~Water Use Reduction, 30% Reduction~~	
ENERGY & ATMOSPHERE		
Credit 1.1	Optimize Energy Performance, 10.5% New or 3.5% Existing	
Credit 1.2	Optimize Energy Performance, 14% New or 7% Existing	EA1 LOW
Credit 1.3	Optimize Energy Performance, 17.5% New or 10.5% Existing	
Credit 1.4	Optimize Energy Performance, 21% New or 14% Existing	EA1 MIDDLE
Credit 1.5	Optimize Energy Performance, 24.5% New or 17.5% Existing	
Credit 1.6	Optimize Energy Performance, 28% New or 21% Existing	
Credit 1.7	Optimize Energy Performance, 31.5% New or 24.5% Existing	
Credit 1.8	Optimize Energy Performance, 35% New or 28% Existing	EA1 HIGH
Credit 1.9	Optimize Energy Performance, 38.5% New or 31.5% Existing	
Credit 1.10	Optimize Energy Performance, 42% New or 35% Existing	
Credit 2.1	On-Site Renewable Energy, 2.5%	
~~Credit 2.2~~	~~On-Site Renewable Energy, 7.5%~~	
~~Credit 2.3~~	~~On-Site Renewable Energy, 12.5%~~	
Credit 3	Enhanced Commissioning	
Credit 4	Enhanced Refrigerant Management	
Credit 5	Measurement & Verification	
Credit 6	Green Power	

Materials and Resources

Credit 1.1 Building Reuse, Maintain 75% of Existing Walls, Floors & Roof
~~Credit 1.2 Building Reuse, Maintain 100% of Existing Walls, Floors & Roof~~
Credit 1.3 Building Reuse, Maintain 50% of Interior Non-Structural Elements
Credit 2.1 Construction Waste Management, Divert 50% from Disposal
~~Credit 2.2 Construction Waste Management, Divert 75% from Disposal~~
Credit 3.1 Materials Reuse, 5%
~~Credit 3.2 Materials Reuse, 10%~~
Credit 4.1 Recycled Content, 10% (post-consumer + ½ pre-consumer)
~~Credit 4.2 Recycled Content, 20% (post-consumer + ½ pre-consumer)~~
Credit 5.1 Regional Materials, 10% Extracted, Processed & Manufactured Regionally
~~Credit 5.2 Regional Materials, 20% Extracted, Processed & Manufactured Regionally~~
Credit 6 Rapidly Renewable Materials
Credit 7 Certified Wood

Indoor Environmental Quality

Credit 1 Outdoor Air Delivery Monitoring
Credit 2 Increased Ventilation
Credit 3.1 Construction IAQ Management Plan, During Construction
Credit 3.2 Construction IAQ Management Plan, Before Occupancy
Credit 4.1 Low-Emitting Materials, Adhesives & Sealants
Credit 4.2 Low-Emitting Materials, Paints & Coatings
Credit 4.3 Low-Emitting Materials, Carpet Systems
Credit 4.4 Low-Emitting Materials, Composite Wood & Agrifiber Products
Credit 5 Indoor Chemical & Pollutant Source Control
Credit 6.1 Controllability of Systems, Lighting
Credit 6.2 Controllability of Systems, Thermal Comfort
Credit 7.1 Thermal Comfort, Design
Credit 7.2 Thermal Comfort, Verification
Credit 8.1 Daylight & Views, Daylight 75% of Spaces
~~Credit 8.2 Daylight & Views, Views for 90% of Spaces~~

Innovation and Design

~~Credit 1.1 Innovation in Design: Provide Specific Title~~
~~Credit 1.2 Innovation in Design: Provide Specific Title~~
~~Credit 1.3 Innovation in Design: Provide Specific Title~~
~~Credit 1.4 Innovation in Design: Provide Specific Title~~
~~Credit 2 LEED® Accredited Professional~~

Appendix C: Results

Table C-1 displays the 641 results of the LEED-NC association rule mining model sorted by importance from highest to lowest. Each rule is categorized as interesting, inexplicable, trivial, or repeat. The recurring synergy groups are listed for reference. The highlighted rules were chosen to be reviewed by the USGBC Technical Development Team. For more details refer to the evaluation subsection in chapter 3.

Table C-1. LEED-NC Association Rule Mining Model Results

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
50%	0.74	EQ 6.1, SS 4.2 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
56%	0.71	SS 5.1, EA 1 High -> EA 2.1	EA 2.1 -> EA 1 High	Repeat
68%	0.68	EQ 7.1, EQ 4.3 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
68%	0.65	EQ 7.1, MR 5.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
51%	0.64	EQ 6.1, EQ 7.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
65%	0.62	EQ 2, EQ 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
52%	0.60	EQ 6.1, SS 7.2 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
51%	0.59	EQ 6.1, EQ 8.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
50%	0.58	EQ 6.1, EQ 3.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
59%	0.58	MR 7, EQ 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
68%	0.57	EQ 7.1, MR 4.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
51%	0.56	EQ 6.1, SS 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
73%	0.56	EQ 6.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Trivial
76%	0.55	EQ 6.2, EQ 4.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
74%	0.55	EQ 6.2, MR 5.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
53%	0.55	EQ 6.1, EA 1 High -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
50%	0.54	EQ 6.1, EQ 4.4 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
74%	0.53	EQ 6.2, SS 4.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
80%	0.53	EQ 6.2, SS 5.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
75%	0.53	EQ 6.2, MR 2.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
73%	0.53	EQ 6.2, MR 4.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
56%	0.53	WE 2, EQ 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
69%	0.53	EQ 7.1, WE 1.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
87%	0.53	EA 2.1, EQ 8.1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
73%	0.52	EQ 6.2, WE 1.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
69%	0.52	EQ 7.1, EQ 4.2 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
74%	0.52	EQ 6.2, WE 3.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
72%	0.52	EQ 6.2, EQ 4.3 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
86%	0.52	EA 2.1, SS 6.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
85%	0.51	EQ 6.2, SS 6.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
91%	0.51	EA 2.1, SS 5.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
73%	0.51	EQ 6.2, EQ 4.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
73%	0.51	EQ 6.2, SS 1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
75%	0.51	EQ 6.2, EQ 5 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
77%	0.50	EA 2.1, EQ 4.2 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
60%	0.50	MR 3.1, WE 3.1 -> WE 2	MR 3.1 -> WE 2	Interesting
74%	0.50	EA 2.1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
75%	0.50	EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Trivial
51%	0.50	SS 5.1, EQ 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
77%	0.50	EA 2.1, WE 1.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
81%	0.50	EA 2.1, SS 7.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
76%	0.49	EA 2.1, WE 3.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
54%	0.49	SS 2, EQ 6.1 -> EQ 6.2	EQ 6.1 -> EQ 6.2	Inexplicable
82%	0.49	EQ 6.2, EA 1 High -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
85%	0.49	EQ 6.2, MR 7 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
75%	0.49	EA 2.1, EQ 4.3 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
61%	0.49	MR 3.1, EA 4 -> WE 2	MR 3.1 -> WE 2	Inexplicable
76%	0.49	EQ 6.2, EQ 8.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
76%	0.49	EQ 6.2, SS 7.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
79%	0.49	EQ 6.2, EQ 4.4 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
75%	0.49	EA 2.1, EQ 4.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
76%	0.48	EA 2.1, MR 2.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
61%	0.48	MR 3.1, SS 6.2 -> WE 2	MR 3.1 -> WE 2	Inexplicable
73%	0.48	EQ 6.2, SS 4.4 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
60%	0.48	MR 3.1, EA 1 High -> WE 2	MR 3.1 -> WE 2	Inexplicable
59%	0.48	MR 3.1, SS 6.1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
58%	0.48	MR 3.1, SS 7.2 -> WE 2	MR 3.1 -> WE 2	Inexplicable
85%	0.48	EA 2.1, WE 2 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
82%	0.48	EA 2.1, EQ 6.1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
72%	0.47	EQ 6.2, SS 4.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
74%	0.47	EA 2.1, SS 4.2 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
80%	0.47	EA 2.1, EA 6 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
73%	0.47	EA 2.1, MR 4.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
73%	0.47	EA 2.1, MR 5.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
88%	0.47	EA 2.1, EQ 6.2 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
76%	0.47	EA 2.1, SS 4.4 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
70%	0.47	EQ 6.2, EQ 7.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
86%	0.47	EA 2.1, EQ 2 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
77%	0.47	EQ 6.2, SS 8 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
69%	0.47	EQ 7.1, EQ 4.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
73%	0.46	EQ 6.2, SS 7.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
73%	0.46	EQ 6.2, EQ 3.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
79%	0.46	EA 2.1, SS 6.2 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
77%	0.46	EA 2.1, EA 4 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
77%	0.46	EA 2.1, SS 5.2 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
81%	0.46	EA 2.1, SS 4.3 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
74%	0.46	EA 2.1, EQ 7.1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
56%	0.46	MR 3.1, EQ 4.4 -> WE 2	MR 3.1 -> WE 2	Inexplicable
79%	0.46	EA 2.1, EQ 4.4 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
74%	0.46	EQ 6.2, SS 6.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
84%	0.45	EA 2.1, EQ 6.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
78%	0.45	EQ 1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
76%	0.45	EA 2.1, EQ 3.2 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
79%	0.45	EQ 6.2, WE 2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
75%	0.45	EA 2.1, EQ 3.1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
71%	0.45	EA 2.1, SS 1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
62%	0.45	EQ 6.2, SS 5.1 -> EQ 2	EQ 6.2 -> EQ 2	Interesting
56%	0.45	MR 3.1, SS 7.1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
75%	0.44	EA 2.1, EQ 1 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
70%	0.44	EQ 6.2, EA 4 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
76%	0.44	EA 2.1, SS 8 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
66%	0.44	EQ 7.1, MR 2.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
67%	0.44	EQ 7.1, SS 1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
53%	0.44	MR 3.1, EQ 5 -> WE 2	MR 3.1 -> WE 2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
53%	0.44	MR 3.1, SS 5.2 -> WE 2	MR 3.1 -> WE 2	Inexplicable
74%	0.43	EA 2.1, EA 3 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
72%	0.43	EQ 6.2, EA 6 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
52%	0.43	MR 3.1, EQ 7.1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
75%	0.43	EQ 3.1, EQ 4.3 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
51%	0.43	MR 3.1, WE 1.1 -> WE 2	MR 3.1 -> WE 2	Interesting
72%	0.42	EA 2.1, SS 7.2 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
71%	0.42	EA 2.1, EQ 5 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
50%	0.42	MR 3.1, EQ 4.2 -> WE 2	MR 3.1 -> WE 2	Inexplicable
50%	0.42	MR 3.1, SS 1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
75%	0.42	EQ 3.1, MR 5.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
50%	0.42	MR 3.1, MR 2.1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
67%	0.41	EQ 7.1, SS 4.2 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
72%	0.41	EQ 6.2, SS 5.1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
78%	0.41	EQ 3.1, WE 1.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
54%	0.41	EQ 2, SS 5.2 -> SS 5.1	SS 5.2 -> SS 5.1	Interesting
69%	0.41	EQ 7.1, WE 3.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
50%	0.40	MR 3.1, EQ 1 -> WE 2	MR 3.1 -> WE 2	Inexplicable
71%	0.40	EQ 7.1, EQ 5 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
70%	0.40	EQ 6.2, SS 4.3 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
50%	0.40	MR 3.1, EA 3 -> WE 2	MR 3.1 -> WE 2	Inexplicable
69%	0.40	EA 2.1, SS 4.1 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
70%	0.40	EA 2.1, EQ 7.2 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
66%	0.40	EQ 6.2, EA 3 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
66%	0.40	EQ 6.2, EQ 3.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
72%	0.39	EA 2.1, MR 7 -> EA 1 High	EA 2.1 -> EA 1 High	Inexplicable
68%	0.39	EQ 6.2, EQ 2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
76%	0.39	EQ 3.1, MR 4.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
72%	0.39	EQ 3.1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
64%	0.39	EQ 6.2, EQ 1 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
56%	0.38	SS 3, SS 6.1 -> EQ 2	SS 3, SS 6.1 -> EQ 2	Inexplicable
55%	0.38	EA 2.1, SS 5.2 -> SS 5.1	SS 5.2 -> SS 5.1	Interesting
71%	0.38	MR 3.1, EQ 6.1 -> EA 1 High	MR 3.1 -> EA 1 High	Interesting
53%	0.38	EQ 6.2, SS 8 -> EQ 2	EQ 6.2 -> EQ 2	Inexplicable
76%	0.37	EQ 3.1, EQ 4.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
69%	0.37	MR 3.1, SS 7.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
55%	0.37	EA 2.1, SS 6.1 -> SS 5.1	EA 2.1, SS * -> SS 5.1	Interesting
68%	0.37	EQ 6.2, SS 2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
55%	0.37	EA 2.1, SS 6.2 -> SS 5.1	EA 2.1, SS * -> SS 5.1	Interesting
70%	0.37	EA 2.1, EA 5 -> EA 1 High	EA 2.1 -> EA 1 High	Interesting
51%	0.37	EQ 6.2, EA 3 -> EQ 2	EQ 6.2 -> EQ 2	Interesting
85%	0.37	EQ 3.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Repeat
74%	0.36	EQ 3.2, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
67%	0.36	MR 3.1, SS 6.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Interesting
67%	0.36	MR 3.1, SS 7.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
50%	0.36	EQ 6.2, EQ 3.2 -> EQ 2	EQ 6.2 -> EQ 2	Interesting
53%	0.36	EQ 6.2, EQ 2 -> SS 5.1	EQ 6.*, EQ 2 -> SS 5.1	Interesting
67%	0.36	MR 3.1, SS 6.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
68%	0.36	MR 3.1, WE 2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
64%	0.35	EA 2.1, SS 8 -> EQ 6.1	EA 2.1 -> EQ 6.1	Interesting
75%	0.35	EQ 3.1, MR 2.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
66%	0.35	EQ 6.2, EA 1 Low -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
52%	0.35	EQ 2, EQ 6.1 -> SS 5.1	EQ 6.*, EQ 2 -> SS 5.1	Interesting
74%	0.35	EQ 3.1, EQ 4.2 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
53%	0.35	EA 2.1, SS 8 -> SS 5.1	EA 2.1, SS * -> SS 5.1	Interesting
66%	0.35	MR 3.1, EQ 7.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
53%	0.35	EA 2.1, SS 4.3 -> SS 5.1	EA 2.1, SS * -> SS 5.1	Interesting
60%	0.34	EQ 6.2, EQ 7.2 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
51%	0.34	EQ 6.2, EA 5 -> EQ 2	EQ 6.2 -> EQ 2	Interesting
75%	0.34	EQ 3.1, SS 1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
65%	0.34	MR 6, SS 5.2 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
63%	0.34	EQ 6.2, EA 5 -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
95%	0.34	EQ 7.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Repeat
64%	0.33	EQ 6.2, EA 1 Middle -> EQ 6.1	EQ 6.2 -> EQ 6.1	Inexplicable
76%	0.33	EQ 3.1, SS 4.2 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
85%	0.33	EQ 3.2, EQ 4.3 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
63%	0.33	MR 3.1, EA 4 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
63%	0.33	MR 3.1, EA 1 High -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
64%	0.33	MR 3.1, SS 6.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
59%	0.33	EA 2.1, SS 5.2 -> EQ 6.1	SS 5.2 -> EQ 6.1	Inexplicable
54%	0.33	EA 1 High, SS 5.2 -> EQ 6.1	EA 1 High -> EQ 6.1	Inexplicable
72%	0.33	EA 3, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
85%	0.33	EQ 3.2, MR 5.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
63%	0.33	MR 6, SS 7.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
71%	0.33	EA 4, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
79%	0.32	EQ 3.1, EQ 5 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
60%	0.32	MR 3.1, SS 5.2 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
57%	0.32	MR 6, SS 7.1 -> MR 7		Interesting
62%	0.32	MR 6, SS 7.2 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
61%	0.32	MR 6, EQ 4.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
95%	0.31	EQ 7.2, EQ 4.3 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
60%	0.31	MR 3.1, SS 5.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
61%	0.31	MR 3.1, SS 8 -> EA 1 High	MR 3.1 -> EA 1 High	Interesting
60%	0.31	MR 3.1, WE 3.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
50%	0.31	EA 1 Middle, SS 4.3 -> EA 5		Inexplicable
95%	0.31	EQ 7.2, MR 5.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
86%	0.31	EQ 3.2, WE 1.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
68%	0.31	SS 6.1, EA 4 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
69%	0.31	SS 4.4, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
60%	0.31	MR 3.1, EQ 7.2 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
88%	0.31	MR 6, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
85%	0.31	EQ 3.2, MR 4.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
58%	0.30	MR 3.1, WE 1.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
86%	0.30	EQ 3.2, EQ 4.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
59%	0.30	MR 3.1, EQ 4.4 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
59%	0.30	MR 6, SS 4.2 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
57%	0.30	WE 2, EQ 6.1 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
58%	0.30	EA 2.1, SS 5.1 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
85%	0.30	MR 3.1, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
59%	0.30	MR 6, SS 5.2 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
95%	0.30	EQ 7.2, MR 4.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
59%	0.30	MR 3.1, EQ 4.4 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
56%	0.30	EA 2.1, SS 6.1 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
81%	0.30	EA 2.1, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
85%	0.30	EQ 3.2, EQ 4.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
58%	0.30	MR 3.1, EQ 5 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
59%	0.30	SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Trivial

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
62%	0.29	SS 6.1, SS 4.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
84%	0.29	MR 3.1, EQ 6.1 -> SS 6.1	MR 3.1 -> SS 6.1	Inexplicable
95%	0.29	EQ 7.2, WE 1.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
58%	0.29	MR 3.1, EQ 3.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
55%	0.29	MR 7, EA 1 High -> EQ 6.1	EA 1 High -> EQ 6.1	Interesting
70%	0.29	SS 5.2, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
60%	0.29	SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Repeat
57%	0.29	MR 3.1, EQ 4.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
60%	0.29	SS 6.1, EQ 4.3 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
81%	0.29	EA 2.1, SS 5.1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
85%	0.29	EQ 3.2, MR 2.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
60%	0.29	SS 6.1, MR 5.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
74%	0.29	EQ 3.1, WE 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
74%	0.29	EA 1 High, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
54%	0.29	WE 2, EQ 7.2 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
57%	0.28	MR 6, WE 1.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
85%	0.28	EQ 3.2, SS 1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
55%	0.28	EA 2.1, SS 6.2 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
77%	0.28	EA 5, EQ 6.1 -> SS 6.2	SS 5.1 -> SS 6.2	Inexplicable
51%	0.28	SS 2, EQ 4.4 -> MR 7		Interesting
56%	0.28	MR 3.1, EQ 7.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
80%	0.28	EA 4, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
61%	0.28	SS 6.2, EQ 4.3 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
83%	0.28	MR 6, EQ 3.2 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
57%	0.28	MR 6, EQ 5 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
96%	0.28	EQ 7.2, MR 2.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
56%	0.28	MR 3.1, EQ 5 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
94%	0.28	EQ 7.2, EQ 4.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
52%	0.28	SS 4.3, EQ 8.1 -> EA 1 High		Inexplicable
55%	0.28	MR 3.1, EQ 4.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
60%	0.28	SS 6.1, WE 1.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
66%	0.28	SS 6.2, SS 4.4 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
66%	0.28	SS 6.2, SS 5.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
95%	0.28	EQ 7.2, SS 1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
56%	0.28	EQ 6.2, WE 2 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
56%	0.28	MR 3.1, SS 7.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
56%	0.28	EA 5, EA 6 -> SS 4.3	EA 5 -> SS 4.3	Inexplicable
58%	0.28	EA 2.1, SS 5.1 -> SS 4.3		Inexplicable
79%	0.28	EQ 6.2, WE 2 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
66%	0.28	SS 6.2, EA 4 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
66%	0.28	SS 4.1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
56%	0.27	MR 3.1, SS 8 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
60%	0.27	SS 6.2, MR 5.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
85%	0.27	EQ 3.2, SS 4.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
55%	0.27	MR 3.1, EQ 3.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
55%	0.27	MR 3.1, EA 4 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
78%	0.27	EA 5, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
54%	0.27	MR 7, EQ 6.1 -> EA 1 High	EQ 6.1 -> EA 1 High	Interesting
54%	0.27	MR 3.1, EQ 4.3 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
51%	0.27	WE 2, EQ 7.1 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
74%	0.27	WE 2, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
95%	0.27	EQ 7.2, EQ 4.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
72%	0.27	SS 5.1, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
60%	0.27	SS 6.1, EQ 4.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
80%	0.27	MR 6, EQ 3.1 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
65%	0.27	SS 6.2, SS 7.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
54%	0.27	MR 3.1, EQ 7.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
50%	0.27	SS 2, EA 6 -> MR 7		Inexplicable
63%	0.27	SS 6.1, EQ 5 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
54%	0.27	EA 2.1, EQ 4.4 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
67%	0.27	SS 7.1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
75%	0.27	EQ 3.1, EQ 7.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
61%	0.27	SS 6.2, EQ 4.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
62%	0.27	SS 6.2, SS 1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
72%	0.27	EA 5, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
61%	0.27	EQ 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
95%	0.27	EQ 7.2, SS 4.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
50%	0.27	EA 1 High, SS 8 -> EQ 6.1	EA 1 High -> EQ 6.1	Interesting
51%	0.27	WE 2, SS 7.1 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
54%	0.27	MR 3.1, SS 4.4 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
53%	0.27	EQ 6.2, SS 8 -> EA 1 High	EQ 6.2 -> EA 1 High	Interesting
53%	0.27	MR 3.1 -> EA 1 High	MR 3.1 -> EA 1 High	Interesting

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
53%	0.26	MR 3.1, MR 5.1 -> EA 1 High	MR 3.1 -> EA 1 High	Interesting
78%	0.26	EA 3, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
63%	0.26	SS 6.1, SS 5.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
63%	0.26	SS 6.1, EQ 7.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
50%	0.26	MR 6, SS 1 -> MR 7		Inexplicable
96%	0.26	EQ 7.2, WE 3.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
69%	0.26	EA 1 High, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
66%	0.26	SS 7.2, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
60%	0.26	SS 6.2, WE 1.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
73%	0.26	EA 2.1, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
62%	0.26	SS 6.2, EQ 4.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
61%	0.26	SS 6.1, WE 3.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
86%	0.26	EQ 3.2, WE 3.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
50%	0.26	EQ 6.1, SS 6.1 -> EA 1 High	EQ 6.1 -> EA 1 High	Interesting
60%	0.26	SS 6.1, MR 2.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
69%	0.26	EQ 6.1, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
52%	0.26	EQ 2, EQ 6.1 -> EA 1 High	EQ 6.1 -> EA 1 High	Interesting
77%	0.26	EQ 3.1, SS 4.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
52%	0.26	MR 3.1, SS 4.2 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
77%	0.26	EQ 1, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
61%	0.26	EQ 1, EQ 4.3 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable
53%	0.26	MR 6, EQ 4.2 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
53%	0.26	MR 6, MR 5.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
53%	0.26	MR 6, WE 3.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
51%	0.26	EA 2.1, SS 7.1 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
86%	0.26	EQ 3.2, EQ 5 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
72%	0.26	WE 2, SS 5.2 -> SS 6.1	WE 2 -> SS 6.1	Interesting
72%	0.26	SS 8, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
74%	0.26	EA 2.1, SS 5.1 -> SS 6.2	SS 5.1 -> SS 6.2	Interesting
78%	0.26	SS 2, WE 2 -> SS 6.1	WE 2 -> SS 6.1	Interesting
97%	0.25	SS 5.1, EQ 6.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
60%	0.25	SS 6.1, EQ 4.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
60%	0.25	SS 6.2, SS 4.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
96%	0.25	EQ 7.2, EQ 5 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
76%	0.25	EA 2.1, EQ 2 -> SS 6.2		Inexplicable
50%	0.25	WE 2, EA 4 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
76%	0.25	MR 6, SS 5.2 -> SS 6.1	SS 5.2 -> SS 6.1	Interesting
71%	0.25	WE 2, EQ 3.2 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
67%	0.25	SS 5.1, SS 1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
71%	0.25	EA 5, EQ 8.1 -> SS 6.2	EA 5 -> SS 6.2	Inexplicable
52%	0.25	MR 3.1, WE 3.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
54%	0.25	SS 5.1, EA 5 -> SS 4.3	EA 5 -> SS 4.3	Inexplicable
52%	0.25	MR 3.1, SS 4.4 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
74%	0.25	WE 2, SS 4.3 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
74%	0.25	SS 4.3, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
50%	0.25	EA 5, EQ 7.2 -> SS 4.3	EA 5 -> SS 4.3	Inexplicable
70%	0.25	WE 2, SS 7.1 -> SS 6.1	WE 2 -> SS 6.1	Interesting
50%	0.25	EA 2.1, EA 1 High -> EQ 6.1	EA 2.1 -> EQ 6.1	Interesting
52%	0.25	MR 3.1, MR 2.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
52%	0.25	MR 3.1, MR 4.1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
75%	0.25	EA 2.1, WE 2 -> SS 6.1	WE 2 -> SS 6.1	Interesting
62%	0.25	SS 6.1, EQ 3.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
62%	0.25	SS 6.2, EQ 7.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
52%	0.25	MR 6, MR 4.1 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
53%	0.25	MR 3.1, EA 3 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
52%	0.25	MR 3.1, EQ 1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
76%	0.25	EQ 3.1, SS 4.4 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
74%	0.25	MR 3.1, SS 5.2 -> SS 6.1	SS 5.2 -> SS 6.1	Interesting
58%	0.25	SS 6.1, MR 4.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
75%	0.25	MR 3.1, SS 7.1 -> SS 6.1	MR 3.1 -> SS 6.1	Interesting
62%	0.24	SS 6.2, EQ 5 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
54%	0.24	EA 1 Middle, EA 5 -> SS 4.3	EA 5 -> SS 4.3	Interesting
51%	0.24	MR 3.1, WE 1.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
51%	0.24	WE 2, SS 8 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
73%	0.24	WE 2, SS 5.1 -> SS 6.1	WE 2 -> SS 6.1	Interesting
60%	0.24	SS 6.2, MR 2.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
52%	0.24	EA 5, SS 6.1 -> SS 4.3	EA 5 -> SS 4.3	Interesting
60%	0.24	SS 6.2, WE 3.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
95%	0.24	EQ 7.2, EQ 3.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
51%	0.24	MR 3.1, EQ 4.1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
50%	0.24	EA 5, SS 4.4 -> SS 4.3	EA 5 -> SS 4.3	Interesting

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
51%	0.24	EQ 6.2, SS 6.2 -> EA 1 High	EQ 6.2 -> EA 1 High	Interesting
76%	0.24	EA 5, EQ 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
61%	0.24	SS 6.1, SS 7.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
74%	0.24	EA 2.1, SS 5.1 -> EQ 8.1	EA 2.1 -> EQ 8.1	Interesting
51%	0.24	MR 6, MR 5.1 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
51%	0.24	MR 6, EQ 4.3 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
50%	0.24	SS 5.1, EA 1 High -> EQ 6.1	EA 1 High -> EQ 6.1	Inexplicable
71%	0.24	EA 2.1, SS 7.1 -> SS 6.1	EA 2.1 -> SS 6.1	Interesting
71%	0.24	SS 5.1, EA 1 High -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
76%	0.24	SS 7.2, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
51%	0.24	MR 6, EQ 4.2 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
97%	0.24	EQ 7.2, EQ 3.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
50%	0.24	EQ 2, SS 5.1 -> EQ 6.1		Inexplicable
76%	0.24	SS 5.1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
74%	0.24	MR 3.1, EA 1 High -> SS 6.1	MR 3.1 -> SS 6.1	Interesting
74%	0.24	MR 6, SS 7.1 -> SS 6.1	MR 6 -> SS 6.1	Interesting
74%	0.24	EQ 2, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Interesting
76%	0.24	SS 7.1, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
75%	0.24	MR 3.1, EQ 8.1 -> SS 6.1	MR 3.1 -> SS 6.1	Inexplicable
61%	0.24	SS 6.1, SS 4.4 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
51%	0.24	EA 2.1, MR 7 -> EQ 6.1	EA 2.1 -> EQ 6.1	Interesting
71%	0.24	WE 2, SS 5.1 -> SS 6.2	SS 5.1 -> SS 6.2	Interesting
62%	0.24	EQ 1, EQ 4.2 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable
78%	0.24	EQ 7.2, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
51%	0.24	MR 6 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
50%	0.24	MR 3.1, EQ 4.2 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
64%	0.24	SS 6.1, EQ 7.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
51%	0.24	MR 6, EQ 4.3 -> EA 1 High	MR 6 -> EA 1 High	Inexplicable
51%	0.24	WE 2, SS 5.1 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
83%	0.24	MR 6, EQ 3.2 -> EQ 4.4	MR 6 -> EQ 4.4	Interesting
83%	0.24	MR 3.1, SS 4.3 -> EQ 4.4		Inexplicable
64%	0.24	SS 6.2, EQ 3.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
74%	0.24	MR 3.1, WE 2 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
82%	0.24	MR 6, EA 4 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
53%	0.24	EQ 6.2, EA 5 -> SS 4.3	EA 5 -> SS 4.3	Inexplicable
65%	0.24	SS 6.2, EQ 7.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
94%	0.24	EQ 7.2, EQ 1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
59%	0.24	SS 6.2, MR 4.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
72%	0.24	MR 3.1, EQ 3.2 -> SS 8	MR 3.1 -> SS 8	Inexplicable
72%	0.24	MR 3.1, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
87%	0.24	EQ 3.2, EA 4 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
87%	0.24	EQ 3.2, SS 4.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
66%	0.24	WE 2, SS 1 -> SS 6.1	WE 2 -> SS 6.1	Interesting
74%	0.24	MR 6, SS 4.1 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
68%	0.24	EQ 4.4, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
85%	0.24	EQ 3.2, EQ 7.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
71%	0.23	EQ 6.2, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
67%	0.23	EA 2.1, EA 5 -> EA 6	EA 2.1 -> EA 6	Interesting
71%	0.23	MR 3.1, EA 4 -> SS 8	MR 3.1 -> SS 8	Inexplicable
68%	0.23	SS 5.1, SS 7.2 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
50%	0.23	MR 3.1, SS 1 -> EA 1 High	MR 3.1 -> EA 1 High	Inexplicable
63%	0.23	SS 6.2, EQ 1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
60%	0.23	EQ 1, MR 5.1 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable
50%	0.23	MR 6, MR 4.1 -> EQ 6.1	MR 6 -> EQ 6.1	Inexplicable
50%	0.23	MR 6, SS 1 -> EQ 6.1	MR 6 -> EQ 6.1	Interesting
65%	0.23	EQ 1, EQ 3.1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
51%	0.23	EA 5, EA 1 High -> SS 4.3	EA 5 -> SS 4.3	Interesting
63%	0.23	SS 5.1, WE 1.1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
97%	0.23	SS 2, SS 1 -> SS 4.1	SS 2 -> SS 4.1	Interesting
71%	0.23	WE 2, EA 5 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
70%	0.23	WE 2, EQ 6.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
65%	0.23	WE 2, EQ 4.2 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
96%	0.23	EQ 7.2, EA 4 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
50%	0.23	MR 3.1, EQ 1 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable
67%	0.23	EQ 6.1, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
87%	0.23	EQ 3.2, EQ 1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
94%	0.23	SS 3, EA 5 -> EA 3	EA 5 -> EA 3	Inexplicable
61%	0.23	SS 6.1, EA 3 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
73%	0.23	MR 6, EA 3 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
65%	0.23	SS 5.1, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
50%	0.23	EA 2.1, WE 2 -> EQ 6.1	EA 2.1 -> EQ 6.1	Inexplicable
50%	0.23	MR 3.1, EA 3 -> EQ 6.1	MR 3.1 -> EQ 6.1	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
50%	0.23	MR 1.1, SS 5.2 -> EQ 6.1		Inexplicable
64%	0.23	SS 5.1, MR 2.1 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
64%	0.23	SS 5.1, SS 5.2 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
80%	0.23	MR 6, EQ 7.1 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
73%	0.23	EA 2.1, WE 2 -> EQ 8.1	EA 2.1 -> EQ 8.1	Interesting
84%	0.23	SS 5.1, WE 1.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
70%	0.23	EA 2.1, SS 5.2 -> SS 6.1	SS 5.2 -> SS 6.1	Interesting
50%	0.23	EQ 6.2, MR 7 -> EA 1 High	EQ 6.2 -> EA 1 High	Interesting
50%	0.23	WE 2, EQ 2 -> EA 1 High	WE 2 -> EA 1 High	Inexplicable
58%	0.23	SS 5.2, SS 4.4 -> SS 6.1	SS 5.2 -> SS 6.1	Inexplicable
61%	0.23	SS 6.1, EQ 1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
95%	0.23	SS 2, EQ 4.3 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
69%	0.23	EA 2.1, EA 1 High -> EQ 8.1	EA 2.1 -> EQ 8.1	Interesting
61%	0.23	SS 6.2, EQ 3.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
95%	0.23	SS 2 -> SS 4.1	SS 2 -> SS 4.1	Interesting
71%	0.23	EA 2.1, EQ 6.1 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable
76%	0.23	SS 5.2, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
69%	0.23	WE 2, EA 5 -> SS 6.2	WE 2 -> SS 6.2	Interesting
87%	0.23	EQ 3.2, SS 7.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
65%	0.23	SS 5.1, SS 4.4 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
72%	0.23	EA 2.1, EQ 2 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable
83%	0.23	SS 5.1 -> SS 5.2	SS 5.1 -> SS 5.2	Trivial
97%	0.22	SS 2, SS 7.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
69%	0.22	EA 2.1, EQ 6.1 -> SS 8	EA 2.1 -> SS 8	Interesting
79%	0.22	MR 6, SS 7.2 -> EQ 4.4	MR 6 -> EQ 4.4	Interesting
95%	0.22	SS 2, MR 5.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
64%	0.22	EA 2.1, SS 4.3 -> EA 6	EA 2.1 -> EA 6	Interesting
86%	0.22	EA 3, EQ 3.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
65%	0.22	SS 5.1, EA 4 -> SS 6.2	SS 5.1 -> SS 6.2	Inexplicable
62%	0.22	SS 6.2, EA 3 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
90%	0.22	WE 2, EA 5 -> EA 3	EA 5 -> EA 3	Interesting
96%	0.22	EQ 7.2, SS 7.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
83%	0.22	SS 5.1, EQ 4.3 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
63%	0.22	SS 6.2, SS 4.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
63%	0.22	SS 5.1, MR 5.1 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
96%	0.22	SS 2, SS 4.2 -> SS 4.1	SS 2 -> SS 4.1	Interesting

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
63%	0.22	SS 5.1, EQ 5 -> SS 6.2	SS 5.1 -> SS 6.2	Inexplicable
70%	0.22	WE 2, MR 7 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
71%	0.22	MR 6, SS 4.2 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
67%	0.22	WE 2, EQ 3.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
96%	0.22	SS 2, EQ 4.2 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
65%	0.22	WE 2, EQ 5 -> SS 6.2	WE 2 -> SS 6.2	Inexplicable
51%	0.22	EA 5, EQ 6.1 -> SS 4.3	EA 5 -> SS 4.3	Inexplicable
66%	0.22	WE 2, EQ 3.2 -> SS 6.2	WE 2 -> SS 6.2	Inexplicable
62%	0.22	SS 5.1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
65%	0.22	WE 2, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
85%	0.22	EA 5, SS 6.1 -> EA 3	EA 5 -> EA 3	Inexplicable
67%	0.22	SS 5.1, EA 1 High -> SS 6.2	SS 5.1 -> SS 6.2	Interesting
65%	0.22	WE 2, EQ 4.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
66%	0.22	WE 2, EA 4 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
96%	0.22	SS 2, MR 2.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
64%	0.22	SS 6.2, EQ 8.1 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
81%	0.22	SS 3, EA 5 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
68%	0.22	WE 2, EA 1 High -> SS 6.1	WE 2 -> SS 6.1	Interesting
83%	0.22	SS 5.1, MR 5.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
80%	0.22	MR 6, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
66%	0.22	WE 2, EQ 5 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
95%	0.22	EQ 7.2, EA 3 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
64%	0.22	EQ 6.1, EA 3 -> SS 6.2		Inexplicable
57%	0.22	SS 6.1, SS 1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
51%	0.22	SS 3, EA 6 -> SS 4.3		Interesting
70%	0.22	SS 5.1, MR 7 -> EQ 8.1		Interesting
63%	0.22	EQ 6.1, SS 7.1 -> SS 6.2		Inexplicable
95%	0.22	SS 2, WE 1.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
88%	0.22	EQ 7.2, EQ 3.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
62%	0.22	SS 6.1, EQ 4.4 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
70%	0.22	MR 3.1, EQ 5 -> SS 6.1	MR 3.1 -> SS 6.1	Inexplicable
94%	0.22	SS 2, MR 4.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
94%	0.22	EQ 7.2, SS 4.4 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
72%	0.22	EA 5, EQ 3.1 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
64%	0.22	WE 2, SS 4.2 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
69%	0.22	MR 3.1, EQ 7.2 -> SS 8	MR 3.1 -> SS 8	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
69%	0.22	MR 6, EQ 7.1 -> SS 8		Inexplicable
69%	0.22	EA 2.1, SS 6.1 -> EQ 8.1	EA 2.1 -> EQ 8.1	Interesting
78%	0.22	SS 2, EA 5 -> EQ 7.2	EA 5 -> EQ 7.2	Inexplicable
80%	0.22	MR 6, EQ 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable
77%	0.21	EQ 2, EA 5 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
85%	0.21	EQ 3.2, SS 7.1 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
51%	0.21	EQ 4.2 -> EQ 4.4		Trivial
87%	0.21	SS 5.1, SS 6.2 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
63%	0.21	EA 2.1, SS 5.1 -> EA 6	EA 2.1 -> EA 6	Inexplicable
69%	0.21	EA 2.1, SS 6.2 -> EQ 8.1	EA 2.1 -> EQ 8.1	Interesting
95%	0.21	EQ 7.2, SS 5.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
84%	0.21	EQ 3.2, SS 4.4 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
65%	0.21	SS 5.1, EA 3 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
60%	0.21	SS 6.1, EQ 3.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
60%	0.21	EQ 1, MR 4.1 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable
81%	0.21	EA 5, SS 4.1 -> EA 3	EA 5 -> EA 3	Inexplicable
50%	0.21	EA 2.1, EA 6 -> SS 4.3		Interesting
62%	0.21	SS 5.1, SS 4.2 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
67%	0.21	EA 2.1, SS 5.1 -> SS 8	EA 2.1 -> SS 8	Interesting
65%	0.21	SS 4.3, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
86%	0.21	SS 5.1, SS 6.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
67%	0.21	SS 5.1, EA 6 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
83%	0.21	SS 5.1, EQ 4.2 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
94%	0.21	EQ 7.2, SS 7.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
63%	0.21	EA 5, SS 5.2 -> SS 6.2	EA 5 -> SS 6.2	Interesting
83%	0.21	SS 5.1, MR 4.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
93%	0.21	EQ 7.2, SS 4.1 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
67%	0.21	SS 6.1, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
63%	0.21	SS 5.1, EQ 5 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
61%	0.21	SS 5.1, EQ 4.3 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
77%	0.21	MR 3.1, EQ 6.1 -> EQ 4.4		Inexplicable
96%	0.21	SS 2, EQ 7.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
84%	0.21	SS 5.1, SS 4.4 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
71%	0.21	MR 7, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
70%	0.21	MR 7, EQ 1 -> EQ 4.4		Interesting
61%	0.21	EQ 1, SS 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
67%	0.21	WE 2, SS 4.3 -> SS 6.2	WE 2 -> SS 6.2	Inexplicable
78%	0.21	SS 8, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
61%	0.21	EA 5, MR 2.1 -> SS 6.2	EA 5 -> SS 6.2	Inexplicable
65%	0.21	SS 5.1, EQ 8.1 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
83%	0.21	SS 5.1, MR 2.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
74%	0.21	EA 5, SS 4.3 -> EQ 7.2	EA 5 -> EQ 7.2	Inexplicable
68%	0.21	EQ 6.2, SS 6.2 -> EQ 8.1		Inexplicable
96%	0.21	SS 2, EQ 3.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
83%	0.21	EQ 2, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
67%	0.21	EA 5, EQ 6.1 -> SS 6.1		Inexplicable
67%	0.21	EA 2.1, EA 3 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable
65%	0.21	WE 2, SS 7.2 -> SS 6.1	WE 2 -> SS 6.1	Interesting
62%	0.21	WE 2 -> SS 6.1	WE 2 -> SS 6.1	Interesting
95%	0.21	SS 2, EQ 5 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
96%	0.21	SS 2, EQ 3.2 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
61%	0.21	SS 6.2, SS 7.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
97%	0.21	SS 2, SS 4.4 -> SS 4.1	SS 2 -> SS 4.1	Interesting
82%	0.21	SS 5.1, EQ 4.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
61%	0.21	SS 6.1, EQ 8.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
64%	0.21	WE 2, SS 5.2 -> SS 6.2	WE 2 -> SS 6.2	Interesting
63%	0.21	SS 5.1, EQ 3.1 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
82%	0.21	EA 5, SS 6.2 -> EA 3	EA 5 -> EA 3	Inexplicable
62%	0.21	SS 5.1, WE 3.1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
63%	0.21	WE 2, EQ 4.3 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
72%	0.21	EQ 7.2 -> EQ 1	EQ 1 -> EQ 7.2	Repeat
68%	0.21	MR 6, EQ 4.1 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
95%	0.21	SS 2, SS 7.2 -> SS 4.1	SS 2 -> SS 4.1	Interesting
89%	0.21	SS 5.1, EA 1 High -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
94%	0.21	SS 2, EQ 4.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
64%	0.21	WE 2, EQ 1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
96%	0.20	SS 2, EQ 1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
68%	0.20	EQ 6.2, SS 5.1 -> EQ 8.1		Interesting
77%	0.20	MR 6, SS 8 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
69%	0.20	MR 3.1, EQ 7.2 -> SS 6.1	MR 3.1 -> SS 6.1	Inexplicable
69%	0.20	MR 6, EQ 5 -> SS 6.1	MR 6 -> SS 6.1	Inexplicable
94%	0.20	SS 2, WE 3.1 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
66%	0.20	WE 2, EQ 8.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
65%	0.20	EA 2.1, EA 1 High -> SS 6.1	EA 2.1 -> SS 6.1	Interesting
57%	0.20	EA 4, SS 5.2 -> SS 6.2		Inexplicable
69%	0.20	EA 5, EQ 4.2 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
62%	0.20	EQ 1, EQ 5 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
69%	0.20	MR 3.1, EQ 3.2 -> SS 6.1	MR 3.1 -> SS 6.1	Inexplicable
86%	0.20	SS 2, EA 5 -> EA 3	EA 5 -> EA 3	Interesting
68%	0.20	EA 2.1, SS 4.3 -> EQ 8.1	EA 2.1 -> EQ 8.1	Inexplicable
61%	0.20	SS 8, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
72%	0.20	EA 5, EA 4 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
76%	0.20	MR 6, EA 3 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
71%	0.20	EA 5, EQ 3.2 -> EQ 7.2	EA 5 -> EQ 7.2	Interesting
61%	0.20	EA 1 High, EA 4 -> SS 6.2		Inexplicable
57%	0.20	SS 5.2, SS 7.1 -> SS 6.1	SS 5.2 -> SS 6.1	Interesting
96%	0.20	EQ 4.4, EQ 7.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
64%	0.20	EA 4, EQ 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
64%	0.20	SS 5.1, SS 4.1 -> SS 6.1	SS 5.1 -> SS 6.1	Interesting
64%	0.20	SS 5.1, EQ 7.2 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
61%	0.20	SS 5.1, EQ 7.1 -> SS 6.2	SS 5.1 -> SS 6.2	Inexplicable
79%	0.20	EA 6, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
62%	0.20	SS 4.3, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
61%	0.20	EQ 1, EQ 4.1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
67%	0.20	MR 3.1, SS 7.1 -> SS 6.2		Interesting
86%	0.20	SS 5.1, EQ 8.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
67%	0.20	EA 6, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
76%	0.20	MR 6, EQ 7.2 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
67%	0.20	MR 6, EQ 3.1 -> SS 8		Inexplicable
62%	0.20	WE 2, MR 5.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
100%	0.20	SS 3, SS 2 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
76%	0.20	EQ 7.2, EQ 5 -> EQ 1	EQ 1 -> EQ 7.2	Interesting
67%	0.20	MR 6, SS 6.1 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
86%	0.20	SS 5.1, EQ 4.4 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
66%	0.20	SS 6.2, EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Inexplicable
64%	0.20	EQ 3.2, EQ 1 -> EQ 7.2	EQ 1 -> EQ 7.2	Interesting
87%	0.20	EQ 4.4, EQ 3.2 -> EQ 3.1	EQ 3.1 -> EQ 3.2	Inexplicable
66%	0.20	EA 2.1, EQ 8.1 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
66%	0.20	MR 3.1, EQ 3.1 -> SS 8	MR 3.1 -> SS 8	Inexplicable
72%	0.20	EQ 7.2, EQ 4.3 -> EQ 1	EQ 1 -> EQ 7.2	Interesting
61%	0.20	EQ 6.1, EQ 4.1 -> EQ 8.1		Inexplicable
66%	0.20	SS 5.1, EA 5 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
61%	0.20	SS 6.2, EQ 4.4 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
63%	0.20	SS 5.1, EA 4 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
62%	0.20	WE 2, MR 2.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
64%	0.20	EA 5, SS 6.2 -> SS 6.1	SS 6.1 -> SS 6.2	Inexplicable
61%	0.20	EA 5, EA 3 -> SS 6.2	EA 5 -> SS 6.2	Inexplicable
68%	0.20	MR 6, SS 7.2 -> SS 6.1	MR 6 -> SS 6.1	Interesting
89%	0.20	EQ 6.2, SS 5.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
63%	0.20	WE 2, EA 4 -> SS 6.2	WE 2 -> SS 6.2	Inexplicable
95%	0.20	SS 2, EA 4 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
77%	0.20	SS 6.1, EQ 3.1 -> EQ 3.2	EQ 3.1 -> EQ 3.2	Inexplicable
67%	0.20	EA 2.1, EA 5 -> SS 6.2	EA 5 -> SS 6.2	Interesting
67%	0.20	MR 6, EQ 3.1 -> SS 6.2		Inexplicable
61%	0.20	MR 4.1 -> SS 7.2		Interesting
75%	0.20	MR 6, SS 4.4 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable
94%	0.20	SS 2, EA 5 -> SS 7.1		Inexplicable
98%	0.20	EA 6, EQ 7.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
59%	0.20	EA 2.1, SS 6.1 -> EA 6	EA 2.1 -> EA 6	Inexplicable
66%	0.20	EA 2.1, EQ 1 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable
51%	0.20	SS 5.2 -> SS 6.1	SS 5.2 -> SS 6.1	Interesting
61%	0.20	SS 5.1, MR 4.1 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
60%	0.20	EQ 6.1, MR 4.1 -> EQ 8.1		Interesting
76%	0.20	EQ 7.2, EQ 3.1 -> EQ 1	EQ 1 -> EQ 7.2	Interesting
81%	0.20	SS 5.1, WE 3.1 -> SS 5.2	SS 5.1 -> SS 5.2	Inexplicable
65%	0.20	EA 2.1, EA 4 -> SS 6.1	EA 2.1 -> SS 6.1	Inexplicable
97%	0.20	SS 6.2, EQ 7.2 -> EQ 7.1	EQ 7.1 -> EQ 7.2	Inexplicable
58%	0.20	EA 2.1, EQ 4.1 -> EA 6	EA 2.1 -> EA 6	Inexplicable
60%	0.20	SS 5.1, EQ 4.2 -> SS 6.1	SS 5.1 -> SS 6.1	Inexplicable
59%	0.20	EA 2.1, EQ 3.2 -> EA 6	EA 2.1 -> EA 6	Inexplicable
66%	0.20	EQ 6.2, EQ 2 -> EQ 8.1		Interesting
58%	0.20	SS 6.1, SS 7.2 -> SS 6.2	SS 6.1 -> SS 6.2	Inexplicable
79%	0.20	EA 5, SS 7.2 -> EA 3	EA 5 -> EA 3	Interesting
60%	0.20	EQ 1, WE 1.1 -> EQ 7.2	EQ 1 -> EQ 7.2	Inexplicable

Confidence	Importance	Rule	Recurring Synergies	Comment
100%	0.99	MR 1.3, SS 1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
100%	0.99	MR 1.3 -> MR 1.1	MR 1.3 -> MR 1.1	Trivial
100%	0.98	MR 1.3, MR 5.1 -> MR 1.1	MR 1.3 -> MR 1.1	Inexplicable
67%	0.89	EQ 7.1 -> EQ 7.2	EQ 7.1 -> EQ 7.2	Trivial
100%	0.20	EQ 6.2, SS 2 -> SS 4.1	SS 2 -> SS 4.1	Inexplicable
62%	0.20	SS 5.1, EQ 7.2 -> SS 6.2	SS 5.1 -> SS 6.2	Inexplicable
61%	0.20	WE 2, WE 1.1 -> SS 6.1	WE 2 -> SS 6.1	Interesting
61%	0.20	WE 2, MR 4.1 -> SS 6.1	WE 2 -> SS 6.1	Inexplicable
74%	0.20	MR 6, SS 7.1 -> EQ 4.4	MR 6 -> EQ 4.4	Inexplicable

Appendix D: Input from the USGBC

These credit associations were presented to the USGBC's Technical Development Team on December 17th, 2007 at USGBC headquarters in Washington, DC. After very brief discussion, the list of the following 25 credit associations was left with a member of the team who added the comments shown and submitted them for the study on January 4th, 2008. The following transcript is taken directly from her email:

1. **EA 2.1 -> EA 1 High**
   - Reason:
     - Project priority: energy
     - Benefits of renewables count towards/are compounded in energy savings
   - Implications:
     - Expected and desired

2. **EA 2.1, SS 6.1 -> SS 5.1**
   - Reason:
     - Rural site: open space, more opportunity for bioswales, wind, etc
     - SSc6, 5 more applicable to rural projects
   - Implications:
     - *[left blank]*

3. **EA 2.1, SS 6.2 -> SS 5.1** *[s/a #3]*

4. **EQ 6.2, EQ 3.2-> EQ 2**
   - Reason:
     - Good design team for flexible, effective HVAC design
     - Project priority: IAQ
   - Implications:
     - *[left blank]*

5. **EQ 1 -> EQ 7.2**
   - Reason:
     - IAQ-minded owners
     - EQ7.2 different in NCv2.1 – both EQ1 and 7.2 require monitoring systems

- Implications:
  - Not surprising

6. **EQ 6.2, SS 8 -> EA 1 High**
   - Reason:
     - Good design team
     - 2 hard credits – modeling experience
     - Integrated design: light where you want it, when you want it → energy savings
   - Implications:
     - More experienced teams → better performing buildings (or, documenting compliance with LEED)

7. **EQ 2, EQ 6.1-> EA 1 High**
   - Reason:
     - Integrated design
     - Increased efficacy of natural ventilation (EQ2) compounded with required operable windows (EQ6.1) decrease load on mechanical ventilations system
   - Implications:
     - *[left blank]*

8. **EA 2.1, EA 1 High -> EQ 6.1**
   - Reason:
     - Energy conscious projects use light controllability so users can limit consumption.
     - Project priority: energy efficiency
     - Strong energy/lighting team. Also explains EA1 High, SS8 → EQ6.1
   - Implications:
     - *[left blank]*

9. **MR 6, SS 7.1 -> SS 6.1**
   - Reason:
     - No idea. Anomalous?
   - Implications:
     - *[left blank]*

10. **WE 2, SS 5.1 -> SS 6.2**
    - Reason:
      - On-site water treatment
      - Rural site: importance of reduced sewage load when less infrastructure nearby
    - Implications:
      - *[left blank]*

11. MR 6, EQ 3.2-> EQ 4.4

- Reason:
  - EQ3.2→EQ4.4?? EQ4.4 unlikely consequence
  - Project priority: IAQ
  - Good argument for MR6, EQ4.4 -> EQ3.2
  - (less VOCs installed → more likely to do IAQ testing, easier to meet credit reqs)
- Implications:
  - *[left blank]*

12. SS 5.1, WE 1.1 -> SS 6.1

- Reason:
  - Xeriscape, focus on maintaining adapted, native plants that require no irrigation.
  - Landscaping as stormwater management
- Implications:
  - *[left blank]*

13. EA 2.1, EA 1 High -> EQ 8.1

- Reason:
  - Energy-minded projects use daylight for lighting and heating savings
- Implications:
  - *[left blank]*

14. SS 2 -> SS 4.1

- Reason:
  - Urban location
- Implications:
  - *[left blank]*

15. EA 2.1, SS 4.3-> EA 6

- Reason:
  - Project priority: carbon, energy
  - Green electrical team: familiar with equipment necessary for on-site renewables, recharge stations?

16. WE 2, EA 5 -> EA 3

- Reason:
  - Having technologies (WE2, etc) with metering in place (EA5) makes commissioning easier
  - Project priority: new/advanced technologies
- Implications:
  - Not surprising

  - EA6 most stand-alone of credits.
- Implications:

- *[left blank]*

17. SS 5.1, MR 7 -> EQ 8.1

- Reason:
  - More rural projects? Greater emphasis on connection to outside: open space, wood finishes and daylight
- Implications:
  - *[left blank]*

18. EQ 2, EA 5 -> EQ 7.2

- Reason:
  - Project priority: IAQ
  - EQ 7.2 different in NCv2.1
  - ME team familiar with newer HVAC/metering technology
  - Strong mechanical design with commitment to system optimization
- Implications:
  - Not suprising

19. MR 7, EQ 1-> EQ 4.4

- Reason:
  - Project priority: IAQ
  - Again, EQ4.4 unlikely consequence from project team standpoint
  - Certified wood may also contain chemicals, so MRc7 often at odds with EQc4.4
- Implications:
  - *[left blank]*

20. WE 2 -> SS 6.1

- Reason:
  - Storm water retention systems/rainwater capture
  - Rural site: importance of reduced sewage load when less infrastructure nearby
- Implications:
  - *[left blank]*

21. WE 2, SS 5.2 -> SS 6.2 *[s/a # 10]*

22. MR 6, SS 7.2 -> SS 6.1 *[s/a #9 (anomalous)]*

23. MR 4.1 -> SS 7.2

- Reason:
  - No idea. While the "green roof recycled trays" idea has some merit, the value of recycled content of plastic trays would be slight
- Implications:
  - *[left blank]*

24. SS 5.2 -> SS 6.1

- Reason:
  - Diverting water to onsite wetland and away from storm drains
  - See *[# 2]*
- Implications:
  - *[left blank]*

References

Agrawal, R., Imielinski, T., and Swami, A. (1993) Mining Association Rules Between Sets of Items in Large Databases, *Proceedings of he ACM SIGMOD Conference on Management of Data*: Washington, DC, pp. 207-216.

Agrawal, R., and Skrikant, R. (1994) Fast Algorithm for Mining Association Rules, *Proceedings of the 20th VLDB Conference*: Santiago, Chile, pp.478-479

AF Memo (2001). Sustainable Development Policy. *Memorandum,* Department of the Air Force: USAF/ILE, Washington, DC: 19 December 2001.

AF Memo (2007). Air Force Sustainable Design and Development (SDD) Policy (DRAFT). *Memorandum,* Department of the Air Force: USAF/A7C, Washington, DC: 14 May 2007.

Army Memo (2006). Sustainable Development Policy. *Memorandum,* Department of the Army.

Berry, Michael J.A. and Gordon S. Linoff (2004) Data Mining Techniques, Second Edition. Wiley Publishing, Inc: Indianapolis, Indiana.

Buddhakulsomsiri, J., Siradeghyan, Y., Zakarian, A., & Li, X. (2006). Association rule-generation algorithm for mining automotive warranty data. International Journal of Production Research, 44(14), 2749-2770.

Clementine (2008). <http://www.spss.com/clementine/>.

Chen, H., Lim, E., Foo, S., Khoo, C., Chen, H., Fox, E.,, et al. (2002). From digital library to digital government: A case study in crime data mapping and mining. In E. Lim,, et al. (Eds.), Digital libraries: People, knowledge, and technology. 5th international conference on asian digital libraries, ICADL 2002. proceedings (lecture notes in computer science vol.2555) (pp. 36-52). Germany; Berlin: Springer-Verlag; Thompson-ISI.

City of Portland (2005). *Green City Buildings: Applying the LEED® Rating System,* City of Portland, Oregon: United States

Chen-Fu Chien, & Li-Fei Chen. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications, 34*(1), 280-290.

Cunningham, S. J., & Frank, E. (1999). Market basket analysis of library circulation data. *Neural Information Processing, 1999. Proceedings. ICONIP '99. 6th International Conference on, , 2* 825-830 vol.2.

Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rudiger Wirth (2000) <u>CRISP-DM version 1.0, Step-by-Step Data Mining Guide.</u> www.crisp-dm.org: United States.

Economist. (2007). Business by Numbers. *The Economist,* 15-21 September 2007, V 384, No. 8546, pg 85-87. London, United Kingdom.

Fayyad, Usama M., Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy (1996) <u>Advances in Knowledge Discovery and Data Mining.</u> The AAAI Press: Menlo Park, California.

GSA (2004). *GSA LEED® Cost Study: Final Report,* General Services Administration: United States.

Hand, David, Heikki Mannila and Padhraic Smyth (2001) <u>Principles of Data Mining.</u> The MIT Press: Cambridge, Massachusetts.

Han, Jiawei and Micheline Kambler (2006) <u>Data Mining Concepts and Techniques.</u> Morgan Kaufmann Publishers: San Francisco, California.

Hardin, G. (1968). The Tradgedy of the Commons. *Science,* 13 December 1968, 1243-48.

Katz (2003). *Green Building Costs and Financial Benefits.* Prepared for the State of Massachusetts: United States

Kibert, Charles J. (2005) Sustainable Construction: Green Building Design and Delivery. John Wiley and Sons, Inc: Hoboken, NJ.

Keel (2008). <http://www.keel.es/>.

KDNuggets (2008). <http://www.kdnuggets.com/software/associations.html>.

KDNuggets (2007) <http://www.kdnuggets.com/polls/2007/data_mining_methodology.htm>

Larose, Daniel T. (2005) <u>Discovering Knowledge in Data: An Introduction to Data Mining.</u> John Wiley and Sons, Inc: Hoboken, NJ.

Larose, Daniel T. (2005) Discovering Knowledge in Data: An Introduction to Data <u>Mining.</u> John Wiley and Sons, Inc: Hoboken, NJ.

MOU (2006) Federal Leadership in High Performance and Sustainable Buildings Memorandum of Understanding. *Memorandum of Understanding between 21 US Federal Agencies,* January 2006.

Miner (2008). <http://www-306.ibm.com/software/data/iminer/>.

Matthiess, Lisa F and Peter Morris (2004) Costing Green: A Comprehensive Cost Database and Budgeting Methodology. *Davis Langdon.*

NAVFAC (2005). Sustainable Development Policy. *Memorandum,* Department of the Navy.

OFEE (2006) The Federal Commitment to Green Building: Experiences and Expectations. Office of the Federal Environmental Executive, *http://www.ofee.gov/sb/fgb_report.html.*

Oracle (2008). <http://www.oracle.com/technology/products/bi/odm/index.html>.

Romero, C., Ventura, S., & García, E. (2007) Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, In Press*

USGBC (2008) www.usgbc.org.

USGBC (2005) LEED-NC Green Building Rating System for New Construction and Major Renovations: Version 2.2. *United States Green Building Council*, Washington, DC.

WCED (1987) Our Common Future. *World Commission on Environment and Development,* Oxford University Press: Oxford, UK.

Weka (2008). <http://www.cs.waikato.ac.nz/ml/weka/>.

Zhang, R., & Yang, J. (2006). Association rules based research on man-made mistakes in aviation maintenance: A case study. *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on, , 1* 545-550.

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 074-0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			
1. REPORT DATE *(DD-MM-YYYY)* 23-06-2007	**2. REPORT TYPE** Master's Thesis		**3. DATES COVERED** *(From – To)* August 2006 – March 2008
4. TITLE AND SUBTITLE Mining Association Rules Between Credits in the Leadership in Energy and Environmental Design for New Construction (LEED-NC) Green Building Assessment System			**5a. CONTRACT NUMBER**
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Thomas, Benjamin J., 2Lt, USAF			**5d. PROJECT NUMBER** If funded, enter ENR #
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865			**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GEM/ENV/08-M19
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Gene Mesick (AFCEE) Bldg 728 Brooks City-Base, TX 78235, 240-4216			**10. SPONSOR/MONITOR'S ACRONYM(S)** AFCEE
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.			
13. SUPPLEMENTARY NOTES			

14. ABSTRACT

The Leadership in Energy and Environmental Design (LEED) Building Assessment System is a performance-based tool for determining the environmental impact of a facility from the whole-building perspective. Taking this vision into account, the individual credits that comprise LEED are designed to reward design teams for employing sustainable design strategies that reduce the total environmental impact across several sustainability issues. This study analyzed projects that have been certified in LEED for New Construction (LEED-NC) versions 2.0 and 2.1. Data on the credits achieved by the projects were mined using the Apriori algorithm which produced 641 association rules. These results were then subjectively reduced to the 24 most synergistic credit combinations and were subsequently identified as credit bundles. This study provided insight into credit interplay and its effect on high-scoring sustainable design strategies. Additionally, it shows that no one strategy is systematically employed by sustainable design professionals in the pursuit of LEED certification. This research lays the foundation for the application of data mining techniques to future LEED data sets. Finally, the revealed credit bundles support the assertion that LEED is a tool that rewards whole-building design and reinforces the perception that integrated design teams are a critical element in successful LEED project delivery.

15. SUBJECT TERMS

Leadership in Energy and Environmental Design (LEED), Sustainable Development (SD), data mining, association rule mining, Apriori

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Alfred E. Thal
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565, ext xxxx (alfred.thal@afit.edu)
U	U	U	UU	108	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18